CHARACTERIZATION OF THE VACCINIA VIRUS J3 PROTEIN AS A POSITIVE TRANSCRIPTION ELONGATION FACTOR

By

DONALD RAY LATNER II

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2001

This work is dedicated to all the members of my family who have desired to attend college while lacking the appropriate opportunity. They have indeed been highly educated by life itself and have obtained wisdom that cannot be found in any classroom.

ACKNOWLEDGMENTS

I would like to humbly take this opportunity to affirm before my scientific peers that I believe in God and I acknowledge Him as the creator of our awe inspiring universe. I choose to believe, not from ignorance of modern scientific facts and principles or as an expression of moral elitism, but as an expression of my profound sense of amazement at the realization that "I" exist. The existence of the human body as a conglomeration of water, DNA, RNA, proteins, and fatty acids is astounding, but the coalescence of these molecules in a fashion that provides an environment for our human emotions such as love, joy, sadness, and our sense of wonder is to me, utterly incomprehensible. I am grateful to God for the privilege to catch a glimpse of the beautiful inner workings of His creation that few have ever seen.

I am especially thankful to my parents for so many things including the direction, encouragement, and the mental and financial support that they have generously supplied while I have been in college and graduate school. I have been well blessed just to know them, but to identify them as my parents and to be called their son is truly an honor.

My graduate advisor, Dr. Rich Condit, is wonderful human being. He is not simply an astute scientist, fantastic teacher, keen mentor, and diligent worker, but he is patient, kind, compassionate, and generous. He is therefore a superb role model both as a professional and as a good person. Being a student in his laboratory has been a great delight and a rare privilege. I am certain that I could not have received a better education in any other lab on the planet and I am respectfully grateful for this opportunity.

I acknowledge my colleagues Dr. Cari Lackner, Dr. Ying Xiang, Ms. Jackie Fried, Dr. Nissin Moussatché, Dr. Penni Black, Dr. Susan D'Costa, Mr. Steve Cresawn, and Dr. Cindy Prins for their friendship, patience, and camaraderie throughout my experience in "The Condit Lab." Finally, the the help and guidance of Dr. Art Edison, Dr. Bert Flanegan, Dr. Sue Moyer, Dr. Jim Resnick, and Dr. David Bloom as members of my advisory committee has been greatly appreciated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

CHARACTERIZATION OF THE VACCINIA VIRUS J3 PROTEIN AS A POSITIVE TRANSCRIPTION ELONGATION FACTOR

By

Donald Ray Latner II

December 2001

Chairman: Dr. Richard C. Condit
Major Department: Molecular Genetics and Microbiology

Vaccinia virus is the prototypic orthopoxvirus. It encodes approximately two hundred genes on a 194-kB linear double-stranded DNA genome and replicates strictly in the cytoplasm of host cells. By necessity, the virus encodes many of the factors that are required for nucleic acid metabolism, making it a good model system for studying intricate cellular processes such as the regulation of transcription. Although much is known about the regulation of transcription initiation in both eukaryotic organisms and vaccinia virus, relatively little is understood about the regulation of transcript 3' end formation in general. In attempt to identify proteins that influence vaccinia mRNA 3' end formation, two independent genetic selections were utilized. In the first selection, a mutation in the J3 gene was identified as an extragenic suppressor of a temperature sensitive mutation in the A18R transcript release factor. In the second selection, several J3 mutations were identified that cause the virus to be dependent upon the transcription enhancing drug IBT. Biochemical inspection of these J3 mutants revealed that they fail

to produce large proteins at late times during infection because they make short, 3' truncated postreplicative gene transcripts. By analogy with the previously described mutations in the G2 transcription factor, this evidence indicates that J3 is also a postreplicative positive transcription elongation factor.

Interestingly, the J3 protein has two previously described roles in mRNA modification. It is a (nucleoside-2'-o-)methyltransferase that converts the cap-0 structure at the 5' end of mRNAs to the cap-1 form and at the 3' end it serves as the small stimulatory subunit of the virally encoded poly(A) polymerase. Previous work has demonstrated that the (nucleoside-2'-o-)methyltransferase and poly(A) polymerase stimulatory activities are independent functions of the protein. To investigate the relationship between the transcription elongation factor activity and the two other activities, several site directed J3R mutant viruses were constructed and were evaluated with respect to each of the three J3 functions. The analysis has shown that the positive transcription elongation activity of J3R is a third independent activity of the protein that can be genetically segregated from its two other roles in mRNA modification.

CHAPTER 1
TRANSCRIPTION ELONGATION AND TERMINATION

Introduction

Life on earth manifests itself through a wonderfully diverse number of creatures. Some current estimates suggest that there may be nearly ten million different species on the planet (World Resource Institute 2001). Astonishingly, each of these organisms subscribe to the central dogma of biology. That is, all of them physically exist as the biochemical summation of their genetic makeup. Every living thing has its own collection of genes that are exquisitely tuned for expression in precise temporal and spatial patterns. Therefore, gene expression is regulated at many different levels including transcription, mRNA splicing, translation, protein folding and modification, and RNA or protein turnover. However, it is the regulation of transcription that provides, by far, the most power and finesse to the control of gene expression.

The process of transcription would never occur without RNA polymerases, which are the molecular machines that convert sequence information encoded by genes into RNA transcripts. Three different types of RNA polymerases can be found in nature. Most organisms, including bacteria, some viruses, and all eukaryotes, utilize multi-subunit polymerases that specifically transcribe DNA templates. The other two types of RNA polymerases that have been described are the single subunit DNA-dependent

polymerases and the RNA-dependent polymerases that are encoded by some bacteriophage and viruses, neither of which will be discussed here.

The mechanical actions performed by all multi-subunit RNA polymerases during transcription can be conceptually broken down into three main stages which are: initiation, elongation, and termination (von Hippel 1998). During initiation, the subunits of a polymerase are assembled on promoter sequences at the 5' end of a gene. This process can be enhanced or prevented through the interaction of a wide variety of DNA-binding protein factors and specific DNA sequences (Conaway et al. 1998a; Dvir et al. 2001). After a polymerase is assembled at the promoter, it requires stimulation from initiation factors, some of which expend energy in the form of ATP, to unwind the DNA template and begin transcription (Dvir et al. 1996). It has been observed that at the beginning of transcription, the polymerase complex is relatively unstable until it has transcribed approximately nine nucleotides. Thus, at each step of nucleotide addition there is a competition between the dissociation of the complex from the template and formation of the next phosphodiester bond. This process is often referred to as abortive initiation (Conaway et al. 2000a). Escape from the promoter and the transition of the polymerase into the elongation stage requires additional stabilizing protein factors that are thought to increase the rate of nucleotide addition. In the case of eukaryotic pol II, phosphorylation of the second-largest subunit of the polymerase is also required for the transition from initiation to the elongation stage (Conaway et al. 2000a).

During elongation, the polymerase is quite stable, illustrated by the fact that it is resistant to high salt concentrations *in vitro*. However, it is influenced by a variety of sequence elements and protein factors that can either pause, arrest, or terminate

transcription (Conaway et al. 2000a). There are some subtle discrepancies in the literature as to what defines a paused polymerase versus an arrested polymerase. For the sake of consistency, a paused polymerase is defined here as one that has temporarily halted transcription and is able to spontaneously continue elongation without the influence of extrinsic factors. In contrast, an arrested polymerase is defined here as one that has halted transcription and cannot continue elongation without the assistance of additional factors that either cause termination or that assist the arrested complex so that it can continue elongation. It has been observed that the 3' end of a nascent transcript in a paused or in an arrested polymerase complex is often displaced from the catalytic site. This phenomenon apparently results from the backward movement of the polymerase in the 5' upstream direction and has thus been termed backtracking (Shilatifard 1998a; Nudler 1999). A variety of proteins that help the polymerase readthrough a pause or arrest site have been characterized over the last few years and some of these proteins, as discussed later, are thought to help a backtracked polymerase re-align the transcript 3' end in the catalytic site so that it can continue elongation (Wind and Reines 2000). Conceivably, it may be impossible for some arrested polymerase complexes to continue elongation due to the loss of an essential subunit or through destabilization of the complex by nucleic acid secondary structure and other protein factors. The polymerase then has no alternative but to release the transcript and the template in a process called termination (Richardson 1993).

Transcription termination is defined as the irreversible release of the transcript and polymerase complex from the template DNA. As described in more detail below, it is known that for certain polymerases or for certain genes, termination occurs directly in

response to cis-acting signal sequences. However, termination of transcription often occurs over a wide range of sequences downstream of a gene coding region and appears almost random in nature (Proudfoot 1989a). In comparison to the process of initiation, relatively little is understood about the sequences and factors that influence the general process of termination. The primary reason for this, as mentioned later, is related to the technical difficulties that are associated with designing experiments to address questions about post-initiation events in transcription. However, some information about the control of elongation and termination has accumulated through studies of model systems that provide powerful genetic tools to take apart the transcription complex. The most thoroughly examined polymerase systems are found in the bacteria *Escherichia coli* and the budding yeast *Saccharomyces cerevisiae*. The following sections of this chapter provide a summary of what is currently known about the different methods by which bacterial and eukaryotic RNA polymerases carry out elongation and termination of transcription. This summary is intended to generate a context for understanding the regulation of post-initiation events in yet a third model system which is the subject of the remainder of this work. Specifically, the investigation described in the following chapters is focused on the discovery and characterization of the vaccinia virus J3 protein which influences the process of viral transcript 3' end formation.

Regulation of E.coli RNA Polymerase Elongation and Termination

The E.coli RNA polymerase holoenzyme is composed of five subunits which include β, β', α (present in two copies), and σ. The β, β', and two α subunits comprise the core enzyme and are homologous to the core subunits found in the eukaryotic polymerases I, II, and III. Several different σ subunits exist and provide the core enzyme

with the ability to initiate transcription at various specific promoter sequences. Shortly after the initiation of transcription, the core enzyme leaves the σ subunit at the promoter and enters the elongation phase of transcript synthesis (Lodish and Darnell 1995).

Transcription elongation can be described as very dynamic or discontinuous process (Conaway et al. 2000b). As previously mentioned, an actively transcribing polymerase often pauses for a variety of reasons including: nucleotide starvation, physical barriers such as DNA binding proteins, or in response to particular template sequences. Some template sequences are barriers to elongation because they form weak hybrids with the transcript or because they promote secondary structure in the transcript. Additionally, stretches of homopolymeric residues may cause the polymerase to slip during elongation and introduce mis-pairing between the RNA-DNA hybrid, thus leading to a pause (Shilatifard 1998b). Regardless of the mechanism, it has been observed that a paused polymerase often appears to slide backwards (Nudler 1999). When this occurs, the polymerase cannot add more residues because the 3' end of the transcript is dislocated from the catalytic site. *E. coli* encodes two elongation factors, called GreA and GreB, that are important for resetting the transcript 3' end in the catalytic site and can thus help a stalled polymerase continue elongation (Kulish, et al. 2000). GreA and GreB are conserved in the bacterial kingdom and the *E. coli* proteins share 35% identity and 57% similarity at the amino acid level. Both proteins have similar but slightly different effects on a stalled polymerase complex. GreA and GreB both stimulate a backtracked polymerase to cleave the nascent protruding transcript, and thereby restore the 3' end in the catalytic site. However, GreA-induced transcript cleavage produces short RNAs that are 2-3 nucleotides long whereas GreB-induced cleavage produces RNAs that are 2-18

nucleotides long (Kulish et al. 2000). Thus, GreA appears to be more specialized for inducing cleavage by polymerases that have backtracked only short distances and GreB is tuned to induce cleavage by polymerases that have backtracked greater distances. Moreover, GreA is unable to induce cleavage in a preformed arrested complex, unlike GreB. This suggests that GreA must be present in the polymerase complex before it pauses, in order to induce cleavage. By inducing cleavage, both GreA and GreB can reduce the amount of time spent at a pause-site. Less time spent dwelling at a pause-site reduces the chance that the polymerase will become destabilized thus prevents premature termination (Kulish et al. 2000). As discussed in later sections, the eukaryotic polymerases also encode cleavage-inducing factors that have similar activities as GreA and GreB.

There are two key mechanisms of termination in *E.coli*. One mechanism is called factor-independent termination and it requires no other protein factors besides the core polymerase itself to facilitate release of the elongation complex and transcript from the template. The second mechanism is called factor-dependent termination and it requires the action of a protein called rho to induce termination. In addition, two processes called attenuation and antitermination are used to modulate the basic factor-independent and dependent termination mechanisms. Specifically, attenuation leads to premature termination of transcription while antitermination prevents it. Thus, the expression of some bacterial genes can be regulated at the level of termination (Lodish and Darnell 1995; Henkin 1996; Henkin 2000).

Factor-independent termination occurs when the polymerase transcribes a particular sequence that has two characteristic features. First, these termination sites have

a short GC-rich section that can self-anneal and produce a stable hairpin structure. Second, a stretch of 6 to 8 U residues is located just downstream of the hairpin (Lodish and Darnell 1995; Mooney et al. 1998) . The precise mechanism by which this sequence induces termination is not completely understood. One model suggests that the hairpin itself could destabilize the polymerase complex through direct interaction. Alternatively, it could indirectly stimulate termination by disrupting the transcript/template hybrid or by disrupting the interaction between the polymerase and the single stranded RNA that has not yet emerged from the complex (Lodish and Darnell 1995). Some evidence supports a direct interaction between the hairpin and the polymerase, but does not completely rule out all of the possible indirect effects (Mooney et al. 1998). For example, the size of the hairpins are conserved to a 5 to 9 nucleotide long stem with a 3 to 5 nucleotide long loop. In some cases, it has been demonstrated that these hairpins can be crosslinked to the β subunit of the polymerase. Additionally, it is known that factor-independent termination sites are active even when the template is single stranded DNA. This evidence suggests that neither the non-template strand nor the transcription bubble is required for termination. It therefore seems likely that the hairpin could directly disrupt an interaction between the transcript and a RNA binding site on the polymerase, thus destabilizing the complex. The most probable mechanism by which factor-independent termination occurs is that when the polymerase reaches the stretch of U residues, it pauses because of the relatively weak nature of the transcript/template hybrid. Typically, a stalled complex would backtrack along the template, cleave the nascent transcript in a fashion mediated by GreA or GreB, and proceed to elongate. However, in this case, the hairpin could prevent the complex from backtracking. The combination of the hairpin interaction with

a RNA binding site on the polymerase and the unstable nature of the U-rich transcript-template hybrid could sufficiently destabilize the complex so that termination could occur. Evidence supporting this model is provided by the observation that a terminator site can be converted to a pause site by simply increasing the distance between the hairpin and the U-rich stretch with a few additional nucleotides. Those extra nucleotides would presumably allow the polymerase enough room to backtrack, cleave, and proceed (Mooney et al. 1998).

As mentioned above, the termination mechanisms in *E.coli* can be modulated to alter gene expression depending on the requirements of the cell. Factor-independent termination can be modified by a system called attenuation that plays a role in the expression of some bacterial genes including the operons that are required for amino acid biosynthesis (Lodish and Darnell 1995). As an example, the tryptophan (*trp*) operon will be considered here. Under conditions of relatively high tryptophan concentration, the amino acid binds to a repressor protein that in turn binds to the *trp* operator sequence and prevents the initiation of transcription. However, some transcription still occurs even in the presence of tryptophan. To prevent those polymerases that actually initiate from proceeding to transcribe the entire operon, a factor-independent terminator hairpin causes the premature termination of transcription at a site located approximately 140 nucleotides downstream of the initiation site. Interestingly, there are two sequence elements upstream of the hairpin that modulate its efficiency as a terminator under varying concentrations of tryptophan. The 5' most sequence element contains two successive tryptophan codons. The second element is located between the tryptophan codons and the terminator hairpin. This second element is capable of hybridizing to the bases located

in the 5' half of the terminator hairpin stem and can thus prevent the formation of the terminator hairpin. As the transcript is being synthesized, a ribosome begins translating it. Under high concentrations tryptophan, the ribosome can readily translate through the region containing the tryptophan codons and cover the second sequence element. With the second element obscured by the ribosome, nothing prevents the formation of the terminator hairpin and premature termination occurs. Conversely, under low concentrations of tryptophan, the ribosome pauses at the tryptophan codons in the 5' element to search for *trp*-tRNAs. While the ribosome is paused, transcription continues downstream and the second sequence element is synthesized. As the 5' half of the terminator hairpin stem emerges from the polymerase complex, the second sequence element binds to it and prevents the terminator hairpin from forming. The polymerase is then free to continue transcribing the remainder of the operon (Lodish and Darnell 1995; Korzheva and Mustaev 2001). This process of attenuation thus serves as yet another way to fine tune gene expression based upon the metabolic needs of the cell.

Factor-dependent termination is the other major way that termination occurs in *E.coli*. Approximately 50% of all terminator sequences in *E.coli* require the rho protein (Lesnik et al. 2001). Rho-dependent termination was initially discovered during in vitro transcription studies of λ-phage DNA using purified *E.coli* RNA polymerase (Roberts 1969; Roberts 1975; Roberts 1988). The DNA template used for the transcription reactions contained two λ promoters called P_L and P_R. The purified polymerase synthesized transcripts from the promoters that were several thousand nucleotides long. It was known, however, that soon after a λ infection, the transcripts synthesized in vivo from these promoters were much shorter and were discrete in length; that is about 1000

and 500 nucleotides for P_L and P_R respectively. When extracts from uninfected cells were added to the purified polymerase, the transcripts produced in vitro matched the length of the transcripts produced in vivo. It was thus implied that there was a component of the uninfected cell extract that caused the polymerase to terminate at the discrete sites, called T_L and T_R. The termination factor was subsequently purified from uninfected cell extracts and is now known as the rho protein (Roberts 1969; Roberts 1975; Roberts 1988). Interestingly, the 50-kDa rho protein shares 26% amino acid identity and 58% similarity with the F_1 ATP synthase that is found in mitochondria (Stitt 2001). Although the rho crystal structure has not been solved, it is known that the protein forms a hexamer that is composed of three homodimers and is therefore structurally very similar to F_1. In addition, the hexamer has three catalytic sites that can bind and hydrolyze ATP in a RNA dependent fashion and it is able to wrap approximately 70-80 nucleotides of a nascent transcript around itself (Stitt 2001). Rho binds to sequences in transcripts that are termed *rut* sites (for rho utilization). Many of these *rut* sites have been identified in *E.coli*, but a comparison of the sequences has revealed no apparent consensus with the exception of being very C-rich (Graham and Richardson 1998). The precise mechanism by which rho induces termination is not clearly understood. However, evidence suggests that the rho dimers are assembled around a nascent transcript so that the RNA is actually threaded through the middle of the hexamer and is additionally wrapped around the outside of the hexamer (Burgess and Richardson 2001). The sequential hydrolysis of ATP bound in the three catalytic sites then drives the 5'-3' translocation of rho along the transcript (Stitt 2001). It has been hypothesized that rho travels down the RNA until it catches up to a polymerase complex that may be paused or

stalled. Upon reaching the polymerase, it is thought that rho facilitates the dissociation of the complex by exerting a 5'-3' RNA-DNA helicase activity (Geiselmann et al. 1993; Kim and Patel 2001). Thus, unlike factor-independent terminators, rho-dependent terminators lack a precisely defined termination point (Richardson and Richardson 1996).

Additional studies of λ-phage revealed yet another system of modulating termination that in contrast to attenuation, prevents the termination of transcription and is hence called antitermination (Lodish and Darnell 1995). In antitermination, the polymerase is converted by protein factors to a termination resistant form that not only reads through downstream terminators, but is less prone to pausing and arrest. Antitermination in λ is induced by two different proteins called N and Q. N prevents the premature termination of transcription from the early promoters P_L and P_R, while Q prevents termination from the late phage promoter (Henkin 1996). Although the mechanism of action for N and Q are slightly different, only N will be considered here.

The N protein is expressed as the first open reading frame from the λP_L promoter and it functions to prevent rho-dependent termination at the T_L and T_R termination sites described above, thereby allowing transcription of downstream genes at the appropriate time during the phage life cycle (Lodish and Darnell 1995). Certain λ mutants have demonstrated that N mediated antitermination requires a sequence element in the nascent transcript, called a *nut* (for N utilization) site, that is located between the terminator and the promoter. *Nut* sites are composed of two sequence elements, called BoxA and BoxB. The BoxA consensus sequence is 12 nucleotides long and is located just upstream of the BoxB element. BoxB forms a hairpin structure composed of a 5 base pair stem (of any complimentary sequence) and a loop that must be 5'-GAAAA-3'. The N protein

specifically binds to the BoxB hairpin via an arginine-rich motif (ARM) that is conserved among many RNA binding proteins. In addition to N, antitermination is coordinated by several other host proteins which have been identified in studies of *E.coli* mutants that are resistant to λ infection. These proteins are referred to as NusA, NusB, NusE, and NusG (Nus stands for N utilization substance). The core antitermination complex is composed of N, NusA, and the *nut* site. This core complex can prevent termination only over short distances away from the *nut* site. However, a complex containing NusG, NusE, and NusB can prevent termination over several kilobases downstream of the *nut* site. As the polymerase elongates away from a *nut* site, the N and Nus proteins remain associated with both the *nut* site and the polymerase. Thus, a loop of RNA is formed. The precise mechanism by which antitermination occurs is not well understood, but several pieces of information have provided hints about the general overall picture. It is known that NusA is a normal component of the polymerase complex that serves to reduce the rate of elongation and it weakly interacts with N. NusB binds to BoxA and is thought to recruit NusE, which is also known as S10 and is a component of the small ribosomal subunit. In addition, NusG has been shown to interact with both the polymerase and with rho. It is hypothesized that N, while bound to BoxB, facilitates assembly of the Nus factors to the polymerase complex and thus stabilizes the polymerase, converting it to a form that is less prone to arrest and that is resistant to the effects of rho. Interestingly, N mediated antitermination can suppress both factor-dependent and factor-independent termination systems. This observation suggests that the stage at which termination is suppressed may be common to both systems. It has therefore been hypothesized that N mediated antitermination could stabilize the polymerase by preventing the disruption of

interactions between: 1) the template and transcript; 2) between the polymerase and transcript; and/or 3) between polymerase and the template (Greenblatt et al. 1993; Lodish and Darnell 1995; Weisberg and Gottesman 1999).

Regulation of RNA Polymerase I Elongation and Termination

In addition to *E.coli*, the next most thoroughly understood termination mechanism is that utilized by eukaryotic RNA polymerase I. Most of the information about pol I termination comes from studies of *S. cerevisiae*, the frog *Xenopus*, and the mouse. Pol I is composed of 14 subunits, including distinct homologues to the core subunits of the *E.coli* enzyme. It is located in the nucleolus where it specifically transcribes the majority of the ribosomal rRNA genes. Additionally, it is the most active polymerase in the cell, accounting for 50-70% of all transcription (Lewin 1994; Lodish and Darnell 1995; Reeder and Lang 1997).

To date, very little information has been published about the factors that regulate the elongation stage of pol I transcription. However, it is known that pol I is served by a cleavage-inducing elongation factor that is functionally analogous to the *E.coli* GreA protein (Labhart 1997; Tschochner 1996). Although the identity of the pol I cleavage factor has not yet been determined, the activity has been detected in both the *Xenopus* and *Saccharomyces cerevisiae* systems. It is known, however, that the factor is not SII, which is the eukaryotic RNA polymerase II cleavage and elongation factor that is discussed in a later section. By analogy with GreA, it has been shown that the pol I cleavage factor induces shortening of the transcript 3' end in a paused polymerase complex by 1-2 nucleotides and thus stimulates the complex to continue elongation (Tschochner 1996). In addition to the cleavage activity, a second pol I elongation factor

called TIF-IC has also been partially purified (Schnapp et al. 1994). TIF-IC has a native molecular mass of 65-kDa, but it is not known if the factor is composed of more than one protein. It has been shown that TIF-IC is required for pol I initiation, and that it stimulates the overall rate of pol I elongation through an unknown mechanism. While it is possible that both the pol I cleavage factor and TIF-IC may function to prevent premature termination, it is not yet known if these two factors have any direct role in the sequence specific pol I termination mechanism that is outlined below.

In contrast to the pol I elongation factors, more is known about the pol I termination mechanism. In fact, pol I termination is relatively straightforward and is thought to be well conserved among eukaryotes (Reeder and Lang 1997). Essentially, pol I termination only requires three components: 1) a special bipartite DNA sequence; 2) a sequence-specific DNA-binding protein; and 3) a transcript release factor. Although more details will be provided below, a brief model of pol I termination suggests that the polymerase stalls when it encounters the terminator protein bound to its specific site in the DNA. The release factor then signals for release through binding the large subunit of the polymerase, the terminator protein, and the 3' end of the nascent transcript. This model is consistent with kinetic studies that have utilized magnetic bead-bound templates which allow the separation of transcripts that have been released through termination from those that are still bound to the template in a paused polymerase complex. These types of experiments have shown that termination occurs in two phases where the polymerase first pauses or arrests then second, the transcript is released. The pausing of the polymerase is mostly controlled by the affinity of the terminator protein to its binding

site in the DNA. Release of the transcript is controlled by the presence of the release factor and the sequence at the 3' end of the transcript (Reeder and Lang 1997).

The DNA sequence of all known pol I terminators is composed of two elements (Reeder and Lang 1997). The first element is an 11-18 nucleotide long binding site for the terminator protein. This binding site is located about 10 bases downstream of the second element which is an A-rich section in the template strand that codes for the last 10-12 nucleotides of the transcript (Lang and Reeder 1995; Mason et al. 1997). The terminator protein binding site facilitates the pausing stage of termination in an orientation specific fashion. The sequences of the binding sites for the terminator proteins found in different organisms do not appear to be conserved. However, the terminator proteins themselves can mediate termination of pol I enzymes from other organisms as long as they are supplied with their cognate binding site at the appropriate location (Kuhn et al. 1990; Reeder and Lang 1997). The A-rich element facilitates the release phase of termination. Although there is no consensus sequence for the A-rich upstream element, replacement of the element with random DNA sequence abolishes release. Briefly, there are two reasons for this. First, this sequence is required in the nascent transcript for efficient binding of the release factor. Second, the sequence is thought to cause limited polymerase slippage that in addition to the release factor, mediates dissociation of the complex. It was hypothesized that replacing the A-rich element with a homopolymeric stretch of A residues would promote more efficient transcript release. However, this was surprisingly not the case (Jeong et al. 1996). When this region is converted to all A residues, the polymerase slips and reiteratively polymerizes thousands of U residues onto the 3' end of the transcript and is unable to

release. The insertion of a single non-homopolymeric residue into this region blocks reiterative polymerase slippage and promotes termination, presumably by preventing complete base-pairing between the RNA-DNA hybrid in the slipped complex. Without a complete hybrid and with the terminator protein blocking further elongation, the complex becomes destabilized and dissociates from the template. As a side note, it appears as if RNA secondary structure plays no obvious role in pol I termination unlike factor-independent termination in *E.coli* (Reeder and Lang 1997).

The *S. cerevisiae* and mouse pol I terminator proteins have been cloned and are known as Reb1p and TTF-1 respectively (Evers et al. 1995; Lang et al. 1994). The ability of Reb1p and TTF-1 to cause polymerase pausing is directly related to the ability of the proteins to bind DNA. Both Reb1p and TTF-1 have an 80 amino acid long DNA-binding domain near the C-terminus that is related to the DNA binding motifs in other transcription regulators such as the Myb oncoprotein, the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR, and the pol III transcription factor TFIIIB. It is thought that the primary function of these proteins is to simply provide a roadblock to pol I elongation that pauses the polymerase over the A-rich section in the terminator sequence (Reeder and Lang 1997). Support for this hypothesis comes from evidence that the non-homologous *E.coli lac* repressor protein and its binding sequence can substitute for the activity of Reb1p and facilitate termination in vitro. However, it was noted that the release efficiency with the *lac* repressor is less than with Reb1p (Jeong et al. 1995). Importantly, these experiments were performed with pol I that was not highly purified. When highly purified pol I was utilized in the *lac* terminator system, release was dependent upon additional yeast protein fractions. This suggested that either Reb1p

directly stimulated the polymerase to release or that there was another unidentified protein which facilitated release in the absence of Reb1p.

Subsequently, a release factor called PTRF (polymerase I transcript release factor) was identified from the mouse by the two-hybrid system using TTF-1 as bait (Jansa et al. 1998). So far, homologues of PTRF have been identified in human and chicken. A search of the yeast genome for a homologous protein has come up empty handed (Jansa et al. 1999a). However, there is evidence indicating that there is a functional homologue in yeast and it has been partially purified. In support of this, PTRF from the mouse can release transcripts synthesized by yeast pol I that is halted by Reb1p. Additionally, PTRF mediates the release of transcripts synthesized by mouse pol I that is stalled by Reb1p but not from yeast pol I stalled by TTF-1. This suggests that some of the molecular interactions required for release are conserved between mouse and yeast, whereas some interactions are species specific. Co-immunoprecipitation experiments have shown that PTRF interacts with both TTF-1 and the large subunit of pol I which is homologous to the β' subunit of *E.coli* RNA polymerase, suggesting that PTRF forms a bridge between the terminator protein and the polymerase. In addition, PTRF binds the 3' end of the nascent transcript in a fashion that is dependent upon the U-rich sequence. Interestingly, PTRF has functional properties that are similar to the La protein which has been implicated as a pol III termination factor as discussed in the next section. Specifically, both PTRF and La mediate transcript release by binding to the 3' end of the nascent transcript. Additionally, both proteins can be isolated in two different forms that are active or inactive for release. In the case of La, the activity is associated with reversible phosphorylation. Unlike factor 2, which is a pol II transcript release factor discussed in a

later section, PTRF mediated transcript release is not ATP-dependent (Jansa et al. 1998; Jansa and Grummt 1999b).

Interestingly, the 3' ends of pol I transcripts are subjected to minor nucleolytic processing where the last 10 or so bases are removed from the transcript by a component of the termination competent polymerase complex. It remains unclear exactly which protein is responsible for this processing event, but it is known that it is not mediated by TTF-1 (Kuhn et al. 1990). By analogy with the TFIIS pol II factor described later, it has been speculated that PTRF could be responsible for the processing event or it may recruit an additional processing factor. With this in mind, the following model has been proposed: After pol I collides with TTF-1, it backtracks or slips revealing the last 3-4 nucleotides of the nascent transcript which are then bound by PTRF. The combination of the slipped mis-paired hybrid and the initiation of 3' processing events destabilizes the polymerase complex so that release can occur (Jansa et al. 1998;Jansa and Grummt 1999b; Jansa et al. 2001).

Regulation of RNA Polymerase III Elongation and Termination

Eukaryotic RNA polymerase III is a 17 subunit enzyme found in the nucleus that is responsible for producing a variety of small transcripts which fall into three different classes based upon the structure and location of their promoters (Geiduschek and Kassavetis 2001). Class I is represented by the 5S rRNA genes. Class II is represented by the tRNAs, 7SL RNA, *Alu* repeat RNAs, the virus associated (VA) transcripts from adenovirus, and some other small viral RNAs. Class III is represented by the small U6 and 7SK RNAs. The promoters of classes I and II are intragenic; that is they are found downstream of the transcription initiation site and are transcribed as the 5' end of the

nascent RNA. Class III promoters are found upstream of the initiation site like pol I and pol II promoters. Several pol III specific transcription factors are required for initiation of each gene class. TFIIIA is specifically required for initiation of class I genes along with TFIIIB and TFIIIC. Initiation of class II genes only requires TFIIIB and TFIIIC, where as class III initiation requires TFIIIB and the factor SNAPc (Gunnery et al. 1999).

Pol III is unique among the RNA polymerases in that most of the transcripts it synthesizes are only about 100-150 nucleotides long, in contrast to pol I and II which transcribe genes that are thousands of nucleotides. In light of this, pol III requires an efficient termination mechanism to rapidly recycle the polymerase for re-initiation of transcription and is therefore, quite prone to termination. Although the mechanism of pol III termination is unclear, some of the fundamental requirements have been suggested (Geiduschek and Kassavetis 2001). It is thought that a small subunit of pol III, called C11, and the second largest subunit, C128, are important for both elongation and termination of transcription (Chedin et al. 1998). C11 is a small essential subunit of yeast pol III that shows homology to some small subunits of human pol I and II as well as the TFIIS (also known as SII) subunit of pol II. By analogy to TFIIS and the *E.coli* elongation factors GreA and GreB, C11 stimulates the pol III intrinsic 3' exonucleolytic activity that shortens a nascent transcript in a paused complex to restore alignment of the active site. As a side note, this pol III nucleolytic activity appears to be an intrinsic property of the enzyme and thus appears more closely related to the intrinsic nucleolytic activity of vaccinia virus RNA polymerase than to the TFIIS-dependent cleavage activity of pol II or the GreA/GreB dependent cleavage of the *E.coli* enzyme (Hagler and Shuman 1993; Geiduschek and Kassavetis 2001). In the absence of C11, pol III is actually less

prone to pausing and is strongly compromised in its ability to efficiently recognize termination signal sequences. Evidence suggests that the C128 subunit is contains the nucleolytic activity of the polymerase and is presumed to bind C11.

It was initially thought that pol III termination simply resulted from the recognition of a terminator sequence by the core polymerase. The terminator sequences from most pol III genes is a short 4-6 base A-stretch in the template strand. However, it was soon realized that the situation was more complicated because there are short A-stretches in the middle of many pol III genes. In addition, a closer inspection of the terminator signals found in several pol III genes revealed that there are important differences between the signal sequences that correlate with gene classification (Gunnery et al. 1999). For example, the 5S rRNA termination signal must be located in a GC-rich sequence context for efficient termination to occur. The substitution of A for GC residues located just upstream of the signal are deleterious to termination. By contrast, the corresponding region just upstream of the termination signals for the U6 and tRNA genes are 4-5 times more A-rich than the 5S rRNA signal. In addition, substitution of A for G residues in the VA RNA_I terminator from adenovirus-2 actually increases termination efficiency. Thus, it seems as if class II and class III transcripts share a similar termination mechanism that is different from the class I genes. It stands to reason that if the pol III holoenzyme itself were sufficient for recognizing the termination signals, all the signals would be the same. Since this is not the case, it has been hypothesized that pol III termination is a factor-dependent process. It is plausible that the differences among promoters of the various gene classes serve as the basis for recruiting different termination factors, and could explain the variability among the sequence

context requirements for the different terminators. For example, TFIIIA is a class I specific initiator and may recruit specific factors that are necessary for termination in a GC-rich context. In contrast, TFIIIC or SNAPc could recruit termination factors that are suited to the more A-rich context of the class II and III terminator sequences (Gunnery et al. 1999; Geiduschek and Kassavetis 2001).

Several proteins have been shown to stimulate pol III termination although a precise role has not been defined for any of them nor is it clear that they are required for termination under all circumstances. Evidence suggests that the TFIIIC initiation factor and some of its associated proteins enhance termination by human but not *S. cerevisiae* pol III (Huang and Maraia 2001; Geiduschek and Kassavetis 2001). It is known that TFIIIC can bind along the length of an entire pol III gene from its promoter through the terminator sequence and is sufficient for release at some terminators. The binding of TFIIIC to the terminator regions of some pol III genes has been shown to be facilitated by the PC4 (positive cofactor 4) and topoisomerase I proteins as well as the heterogeneous NF-1 (nuclear factor 1) family of proteins (Wang and Roeder 1998; Wang et al. 2000). Interestingly, PC4 and topoisomerase I have been previously described as general coactivators of pol II transcription that facilitate the formation of the TFIIA-TFIID complex at the TATA box in the promoter. Additionally, topoisomerase I is important in pol I transcription for relieving the torsional strain associated with elongation. The NF-1 family members bind to the consensus sequence 5'-YTGGCA(N_3)TGCCAR-3' which is found in two different terminators of the pol III VA1 gene. Although it is known that PC4, topo I, and NF-1 proteins enhance pol III termination efficiency through facilitating TFIIIC binding to terminator regions, it is not known if they perform any other specific

function in directly inducing release of the transcript (Wang and Roeder 1998; Wang et al. 2000; Geiduschek and Kassavetis 2001).

It has been reported that the human autoimmune antigen La is an essential pol III termination factor although there has been published evidence to the contrary (Geiduschek and Kassavetis 2001). La is a RNA binding phosphoprotein that has strong affinity for RNAs having UUU-OH-3' as the last three nucleotides and thus commonly associates with pol III transcripts. In addition, it can also bind to the 5'-ppp ends of nascent transcripts and is found associated with the pol III holoenzyme. HeLa extracts depleted of La loose up to 99% of their pol III transcription activity. Additionally, in vitro transcription experiments with immobilized templates have suggested that the unphosphorylated form of human La facilitates pol III recycling by stimulating release and reinitiation of the polymerase from certain terminators (Maraia et al. 1994; Maraia and Intine 2001). However, in contrast to the human La protein, evidence indicates that La from *Xenopus* plays no role in pol III termination at all (Lin-Marq and Clarkson 1998). Thus, the significance of La as a pol III general termination factor remains undetermined.

In summary, a simple model of pol III termination would predict that the pause-release activity of the holoenzyme C11 subunit is switched off by a variety of extrinsic factors that are recruited based upon the sequence context in which the terminator is located. It is possible that polymerase is stalled by either the terminator element itself or by TFIIIC bound to the terminator. With the C11 subunit unable help reset the active site through induction of intrinsic cleavage, the complex is left to dissociate either on its own or by the action of La (Geiduschek and Kassavetis 2001).

Regulation of RNA Polymerase II Elongation and Termination

Eukaryotic pol II is responsible for synthesizing the messenger RNAs (mRNAs) that encode most cellular proteins. The pol II holoenzyme consists of twelve subunits, including core subunits that are homologous to α, β, and β' of the *E.coli* enzyme and six subunits that are shared with pol I and III. In addition to the holoenzyme, a multitude of general and gene-specific factors that are required for pol II initiation have been described. Pol II initiation minimally requires five general initiation factors which are: TFIIB, TFIID, TFIIE, TFIIF, and TFIIH (Conaway et al. 2000b). Interestingly, TFIIF participates in both initiation and elongation. As an initiation factor, TFIIF binds the polymerase and stabilizes its interaction with the initiation factors TFIIB and TFIID. In its role as an elongation factor, TFIIF increases the rate of nucleotide addition and thus reduces abortive initiation and facilitates promoter escape (Conaway et al. 1998b) (Moreland et al. 1999; Yan et al. 1999). Besides TFIIF, several other factors have been described that either positively or negatively affect pol II elongation. As described in more detail below, these factors fall into 4 categories based on their mechanism of action: 1) elongation rate enhancement; 2) transcript cleavage; 3) CTD modification; and 4) chromatin remodeling.

In addition to TFIIF, three other proteins called ELL, elongin, and CSB have been shown to increase the rate of pol II transcription, although the precise mechanism of action for each remains unknown (Conaway et al. 2000b). The ELL elongation factor was identified as a gene that is frequently translocated with the MLL gene in acute myeloid leukemia (AML). It was discovered that ELL is part of a complex called holo-ELL that is formed with three additional proteins called EAP20, EAP30, and EAP45. ELL has

been shown to bind pol II and suppress pausing by stimulating the catalytic rate of the polymerase (Shilatifard 1998c). Elongin is a second factor that stimulates the elongation rate of pol II and is composed of three subunits simply called A, B, and C. It has been shown that elongin is able to increase the catalytic rate of pol II only after the first 8-9 nucleotides have been polymerized, suggesting that pol II undergoes some structural change after leaving the promoter that makes it susceptible to elongin activation. Additionally, elongin has been shown to increase the rate of transcript synthesis by the polymerase regardless of the concentration of free ribonucleotides, suggesting that elongin does not simply decrease the k_{cat} or increase the K_M of the enzyme. It has been hypothesized that during transcription, a polymerase cycles between catalytically active and inactive forms at each step of nucleotide addition. It is thought that elongin functions either by directly converting an inactive polymerase to the active form or by locking an active polymerase into the active conformation until the incoming ribonucleotide is positioned in the active site and the phosphodiester bond is formed (Moreland et al. 1998). Finally, the Cockayne syndrome group B (CSB) protein is thought to play a role in a human inherited developmental disorder that results in "cachectic dwarfism" (Selby and Sancar 1997). Affected patients are mentally retarded due to impaired neurological development and usually die by age 12. Cells from CS patients are defective for transcription-coupled DNA repair, so it was originally hypothesized that CSB was a transcription-repair coupling (TRC) factor. However, evidence now suggests that CSB probably does not have a direct role in transcription-repair, but rather serves to stimulate the rate of pol II elongation. By enhancing elongation, it is thought that CSB increases the rate at which pol II becomes blocked at sites of DNA damage and elicits repair.

Interestingly, CSB has been shown to inhibit pol II cleavage activity induced by a factor called TFIIS (described below) and is thought to directly recruit repair factors to sites of DNA damage. It has been hypothesized that TFIIS may interfere with transcription-repair coupling factors by inducing repeated shortening and elongation of the transcript. Simultaneous inhibition of TFIIS and binding of repair factors would thus facilitate readthrough by an arrested polymerase (Selby and Sancar 1997).

Like the *E.coli* polymerase and the eukaryotic polymerases I and III, pol II also has a pause-suppressing transcript-cleavage factor with an activity that is analogous to GreA and GreB. As just mentioned, the pol II cleavage factor is known as TFIIS (or SII) (Wind and Reines 2000). TFIIS is a zinc-containing 35-kDa protein that stimulates the intrinsic ribonucleolytic activity of pol II to reactivate it from an arrested state. Although it is not known how TFIIS induces cleavage by the polymerase, it is known that it directly binds to the polymerase and requires magnesium for activity. In addition to assisting polymerases that have stalled due to intrinsic template sequences, it has been shown that TFIIS also enables the polymerase to transcribe past some DNA-bound proteins and small DNA-bound drugs. However, it will not enable the polymerase to transcribe past a cyclobutane pyrimidine dimer, which is a common type of UV light induced DNA damage (Wind and Reines 2000).

A third class of transcription elongation factors function by altering the carboxy-terminal domain (CTD) of the second largest subunit of pol II (Conaway et al. 2000b). The pol II CTD is characterized by the heptapeptide YSPTSPS that is repeated between 26 and 52 times from yeast to mammals respectively. Importantly, the CTD can be highly phosphorylated and thus provides a means to regulate various aspects of

transcription. More specifically, the CTD is hypophosphorylated during initiation and becomes hyperphosphorylated during elongation (Hirose and Manley 1998a). Early efforts to identify the role of CTD phosphorylation in transcription utilized a drug called 5,6-dichloro-1-β-D-ribofuranosyl-benzimidazole (DRB), which is a protein kinase inhibitor. Fractionation of DRB-sensitive transcription systems led to the isolation of one DRB-sensitive positive-acting elongation factor called PTEF-b and two negatively acting factors called NELF and DSIF (Conaway et al. 2000b). PTEF-b (positive transcription elongation factor b) is a cyclin dependent kinase composed of the cyclin dependent kinase Cdk9 (also referred to as PITLARE) and one of several cyclins including T1, T2, and K (Price 2000). Current evidence suggests that PTEF-b mediated phosphorylation of the pol II CTD is required to prevent arrest of an elongating complex. In addition, PTEF-b may play a similar role in termination because, as described later, the CTD serves as a type of central signal processing center where many different transcript processing and termination factors can bind depending on the phosphorylation status of the CTD (Price 2000). NELF (negative elongation factor) is a complex that is composed of at least five novel proteins ranging from 46-66 kDa and DSIF (DRB sensitivity-inducing factor) is a heterodimer of 160 and 16 kDa subunits. Interestingly, NELF and DSIF bind to each other and to pol II. Without both NELF and DSIF, purified pol II is not inhibited by DRB. The exact mechanism by which DSIF and NELF modulate elongation is not known. However, it has been shown that they suppress transcription elongation through interactions with a polymerase that contains a hypo-phosphorylated CTD, and are thus antagonized by PTEF-b (Wada et al. 1998; Yamaguchi et al. 1999).

Finally, the fourth class of pol II elongation factors stimulate elongation by modifying chromatin to make it more accessible for the initiation and elongation stages of transcription. Members of this class include the SWI/SNF, HMG-14, FACT, and the elongator protein (Conaway et al. 2000b). SWI/SNF is actually a large family of proteins that remodel nucleosomes and are conserved throughout eukaryotes (Sudarsanam and Winston 2000). These proteins were originally discovered in yeast, where they are important for expression of the HO endonuclease gene that is important for mating type switching (SWItching) and the SUC2 gene which is required for sucrose fermentation (Sucrose Non-Fermenter -SNF). Although the mechanism of action for the SWI/SNF complex is unknown, it has been shown that it changes the rotational phasing of DNA on histone octamers in an ATP-dependent fashion so that the DNase I digestion pattern is altered. Additionally, it has been shown that an ATP-dependent activity in fractions containing the SWI/SNF complex can stimulate pol II elongation from the *HSP70* promoter by remodeling nucleosomes downstream of the transcription start site (Sudarsanam and Winston 2000; Conaway et al. 2000b). In addition to SWI/SNF, HMG-14 is a non-histone protein that preferentially associates with active chromatin and stimulates pol II elongation. It is thought that HMG-14 activates elongation by interacting with the histone protein H1, which normally serves to condense chromatin into a highly condensed, transcriptionally inactive form. Although the mechanism of HMG-14 action is unknown, it appears to disrupt the interactions between adjacent H1 molecules to provide more space between nucleosomes and relax the chromatin structure (Ding et al. 1997; Ding et al. 1994). FACT (facilitates chromatin transcription) is a third protein that stimulates pol II elongation on chromatin templates that specifically lack the

histone H1 protein. Once again, the mechanism of action is not clear, but FACT appears to interact with histone H2A-H2B protein dimers to promote nucleosome disassembly during transcription. The elongator protein has been shown to interact with an elongating pol II that has a hyperphosphorylated CTD and the SII cleavage factor. In addition, it also has histone acetyltransferase activity suggesting that it may stimulate elongation by modifying and destabilizing nucleosomes that are blocking polymerase elongation (LeRoy et al. , 1998; Orphanides et al. 1998).

To summarize, a large number of factors that stimulate pol II elongation have been discovered over the last few years. Although the precise mechanisms by which many of the factors modulate elongation is not known, it seems that there are two major ways that elongation factors work. First, transcript cleavage appears to be a fundamental mechanism by which multi-subunit polymerases, including pol II, overcome pause and arrest sites. This is demonstrated by the fact that all of the polymerases considered thus far have the intrinsic ability to cleave the nascent transcript, and all but pol III have known extrinsic factors that stimulate the cleavage activity when the polymerase becomes paused. The second type of elongation factor appears to directly stimulate the rate of elongation by a number of possible mechanisms. Some of the factors considered here have been shown to directly bind pol II or alter its structural properties. By doing so, these factors could convert or lock the polymerase into an active conformation that is less prone to pausing between each step of nucleotide addition. Finally, the proteins that remodel chromatin have an obvious role in making a condensed template accessible and in clearing the path for an elongating polymerase.

In contrast to pol I and III transcripts, most pol II transcripts are modified in several ways both during and after synthesis. Not surprisingly, pol II termination is thought to be linked to some of these modification steps (Lodish and Darnell 1995; Zhao et al. 1999). Most pol II transcripts are capped and methylated at the 5' end. In addition, primary pol II transcripts have large segments of non-coding sequence called introns that are removed from the final messenger RNA by a process called splicing. The 3' ends of mRNAs are not formed by termination as is the case for transcripts produced by the other polymerases considered thus far. Rather, mRNA 3' ends are formed by a cleavage event that is mediated by several proteins acting together in response to a conserved signal sequence located in the transcript. Cleavage of the transcript frees what will become the mature mRNA from the nascent RNA that is still associated with the pol II enzyme. The mRNA 3' end is polyadenylated while pol II is left to continue elongating the nascent transcript downstream. Pol II then terminates in a seemingly random fashion at various places that may be up to four kilobases downstream of the cleavage/polyadenylation site (Lodish and Darnell 1995). There are only two known classes of pol II transcripts that are not polyadenylated or terminated in this fashion and they will be considered in more detail later. How pol II decides to terminate is unclear. Progress in elucidating the pol II termination mechanism has been hindered because the nascent transcripts, which could perhaps be used to identify termination signal sequences and proteins, are highly sensitive to degradation and are difficult to detect. However, several clues are beginning to provide a framework for understanding the general termination mechanism and are described in more detail below.

It is known that termination of pol II transcription requires a functional polyadenylation signal in addition to certain protein factors that mediate transcript cleavage (Proudfoot 1989b; Lodish and Darnell 1995; Zhao et al. 1999). Moreover, it has been repeatedly observed that the efficiency of termination directly correlates with the strength of the poly(A) signal. The core polyadenylation signal in mammals consists of three sequence elements. First, there is a highly conserved sequence, AAUAAA, that is found 10–30 bases upstream of the cleavage site. Second, there is a less-highly conserved U or GU-rich element that is located within 30 bases downstream of the cleavage site. Third, is the cleavage site itself which eventually becomes the site of poly(A) addition and is essentially determined by the distance between the AAUAAA and the U/GU-rich elements. Often, cleavage occurs on the 3' side of a CA-dinucleotide. In addition to the core signal sequence, multiple protein factors are required for cleavage and polyadenylation (Zhao et al. 1999). Cleavage in mammals requires the action of at least five protein factors including CPSF (cleavage and polyadenylation specificity factor) which is responsible for recognizing the AAUAAA sequence; CstF (cleavage-stimulatory factor) which binds to the U/GU-rich element; CF I_m and CF II_m (cleavage factor I_m and II_m), PAP (the poly(A) polymerase), and pol II itself. Polyadenylation requires PAP, CPSF, and PAB II (poly(A)-binding protein II) (Zhao et al. 1999).

It is thought that the cleavage and polyadenylation machinery associates with pol II at the promoter and is delivered by the polymerase to the 3' end of the transcript. In fact, it has been shown that CPSF binds to the transcription initiation factor TFIID which is important for assembly of the preinitiation complex at pol II promoters (Dantonel et al. 1997). CPSF can be found both in preinitiation complexes and in elongating complexes

formed in vitro, suggesting that the polyadenylation machinery is recruited by TFIID to the preinitiation complex and is then transferred to the polymerase as it begins elongation (Zhao et al. 1999). The association between the cleavage/poly(A) factors and pol II is facilitated by the unique carboxy-terminal domain (CTD) of the largest pol II subunit (Hirose and Manley 1998b). The CTD is characterized by the heptapeptide YSPTSPS that is repeated between 26 and 52 times from yeast to mammals respectively. Importantly, the CTD can be highly phosphorylated and thus provides a means to regulate various aspects of transcription. More specifically, the CTD is hypophosphorylated during initiation and becomes hyperphosphorylated during elongation (Hampsey 1998). It has been shown that CPSF and CstF can bind to both the hypophosphorylated and hyperphosphorylated forms of the CTD (McCracken et al. 1997). Two lines of evidence suggest that CstF and the CTD are required for termination, although it is known that cleavage at the poly(A) site is not a prerequisite for termination (Dye and Proudfoot 2001). First, experiments with yeast that contain temperature sensitive mutations in CF IA, the homologue of mammalian CstF, have shown that cleavage, polyadenylation, and termination are all inhibited at the non-permissive temperature (Birse et al. 1998a). Second, analysis of CTD truncation mutants show that the CTD is required for splicing, cleavage, polyadenylation and termination (McCracken et al. 1997).

Insight to the requirements of the poly(A) signal and the role of various cleavage factors in termination has been provided by a very recent examination of the human β and ε-globin genes by Dye and Proudfoot (2001). They discovered that in addition to poly(A) site cleavage, there is a heterogeneous and reiterative endonucleolytic cleavage

of the nascent transcript that is dependent upon poorly defined AT-rich sequence elements in the 3' flanking region downstream of the poly(A) site. This type of cleavage appears to chop off 5' pieces of the nascent transcript while the polymerase continues elongation. They referred to this phenomenon as "pretermination cleavage" (PTC) and demonstrated that PTC is required for termination from these two genes. It was shown that PTC does not require a functional poly(A) signal and is independent of the processing events at the poly(A) site. In addition, when PTC was abolished by deleting segments of the 3' flanking region, processing at the poly(A) site appeared normal. Furthermore, it was shown that the poly(A) signal is required for release of the transcript and the polymerase from the template. These results are consistent with another recent observation by Osheim *et al.* (Osheim et al. 1999) that suggests termination requires a poly(A) signal, but does not require transcript cleavage at the poly(A) site. Although the protein(s) that mediate PTC cleavage remain unidentified, it is quite possible that some of the same proteins that are required for the poly(A) site cleavage are responsible for PTC as well. Alternatively, PTC could be an intrinsic activity of pol II that is analogous to the activity mediated by the C11 subunit of pol III. In fact, it is known that the pol II elongation factor TFIIS stimulates an intrinsic endonucleolytic activity to help the polymerase through roadblocks during elongation. However, the type of cleavage mediated by TFIIS in a paused complex effectively removes a 3' piece of the transcript, allowing the polymerase to continue elongating the 5' portion. In the case of PTC, it appears as if the 5' upstream piece of the transcript is removed while the polymerase continues elongation of the 3' portion. Despite the fact that there has been no previously described role for TFIIS in termination and while it seems unlikely from a mechanistic

standpoint, TFIIS activity could be modulated in some way to facilitate PTC and termination (Dye and Proudfoot 2001).

Based on these new results, it appears as if pol II termination may occur in two steps (Dye and Proudfoot 2001). Hypothetically, the transcript first undergoes PTC in a poorly-defined sequence-specific fashion. It is predicted that the PTC cleaved pre-mRNA remains associated with the polymerase as it continues elongation of the nascent transcript. This association could facilitate cleavage and polyadenylation at the upstream poly(A) site. The 3' processing machinery, while bound to the poly(A) site, could destabilize the polymerase through direct interaction and induce the second step of termination, which is transcript release. This model is also not inconsistent with a previously described model which proposes that an anti-termination factor is released from the polymerase complex upon transcription of the poly(A) signal (Logan et al. 1987; Proudfoot 1989b). While these exciting new results await independent confirmation, they have offered insight to the role that the poly(A) signal plays in the termination process. However, still unexplained by these results is the observation that termination generally occurs over a broad range of thousands of bases downstream from the poly(A) signal. PTC and 3' processing are probably only one part of the entire process. Indeed, several other factors have been implicated in pol II termination and are described in more detail below (Zhao et al. 1999; Aranda and Proudfoot 2001).

The hallmark of termination is the actual release of the transcript and polymerase from the template. As just mentioned, it is known that the poly(A) signal is required for release, but the mechanism is not understood. A protein called factor 2 was originally characterized as a member of the *Drosophilia* N-TEF (negative transcription elongation

factor) complex and has been shown to facilitate transcript release during premature termination that often occurs upstream of the poly(A) signal (Xie and Price 1997). Although the mechanism of factor 2 mediated transcript release is not known, it has been demonstrated that the protein is a double-stranded DNA dependent ATPase. In addition, factor 2 has seven conserved helicase motifs that are characteristic of the helicase superfamilies 1 and 2. However, it has been shown not to have helicase activity itself (Zhao et al. 1999). Interestingly in this regard, it is similar to the vaccinia virus early gene termination factor NPH-1 (nucleoside phospho-hydrolase 1) that is discussed in the last section of this chapter (Xie and Price 1996). Factor 2 is also a member of the SWI2/SNF2 protein family (Xie and Price 1998). It has been suggested that the SWI2/SNF2 family members use ATP hydrolysis to disrupt protein-DNA interactions. Regardless of its mode of action, the N-TEF complex promotes abortive transcription initiation which is characterized by the reiterative synthesis of short transcripts as the polymerase tries to escape the promoter. The activity of N-TEF is antagonized by the influence of the P-TEF (positive transcription elongation factor) complex. Interestingly, a component P-TEF, called P-TEFb, is a kinase that phosphorylates the CTD of pol II and thus promotes the transition from abortive initiation to processive elongation (Hampsey 1998; Shilatifard 1998a). It could be that the phosphorylation of the CTD disrupts a potential interaction with factor 2 in order to prevent premature termination. Although it has not yet been demonstrated, factor 2 may play a similar role in mediating transcript release during normal termination downstream of the poly(A) site and may require the action of a yet unidentified CTD-specific phosphatase.

Given that there is a great deal of functional conservation between the prokaryotic and eukaryotic transcription systems, it is surprising to note that based on sequence homology, the bacterial termination factor rho does not have a eukaryotic homologue (Richardson 1993; Platt 1994). It is possible, however, that a functional homologue exists and in fact, there is some evidence to support it (Wu and Platt 1993; Lang et al. 1998). Purified yeast pol II can be induced to pause at both intrinsic DNA elements and through the action of DNA binding proteins such as the pol I termination factor Reb1p or the *lac* repressor protein. When pol II is paused in this manner, the addition of bacterial rho protein causes the release of the polymerase in an ATP-dependent fashion if a rho binding site is present in the transcript. Interestingly, rho has no effect on a paused pol I or pol III complex. In conjunction with previous observations, this suggests that the termination mechanism utilized by pol II more closely resembles bacterial rho-dependent termination while the mechanism used by pol I and III appears more similar to rho-independent termination. In addition, several eukaryotic proteins besides factor 2 have helicase motifs and have been shown to play roles in various aspects of transcription from chromatin remodeling and transcription initiation to transcription-coupled repair. By analogy with the rho helicase, it is possible that one or more of these putative helicases plays an important role in pol II termination (Zhao et al. 1999).

As just mentioned, it is known that pol II can be artificially paused by site-specific DNA binding proteins and by sequence elements that intrinsically facilitate pausing. By analogy with pol I, it is conceivable that pausing also plays an important role in pol II termination. One model that is reminiscent of rho-dependent termination in *E.coli* suggests that the 5' unprotected end of the nascent transcript may be subject to

attack by a 5'-3' exonuclease that could chase down the polymerase and disrupt the elongation complex (Eisen and Lucchesi 1998; Zhao et al. 1999). This process could be enhanced if the polymerase were paused at a downstream site and is not necessarily inconsistent with the discovery of pretermination cleavage. In fact, PTC endonucleolytic cleavage could shorten the distance that a hypothetical 5'-3' exonuclease would have to travel to catch up with pol II. Although there does not appear to be any sort of canonical pause element that is required for termination from all pol II genes, a couple of pause sites have been described that may be important for termination from at least two different genes and are worth pointing out. The first example is a 92 basepair element located just downstream of the poly(A) site in the human α2-globin gene which contains six nearly perfect tandem repeats of the pentanucleotide CAAAA within a CG-rich context on the template strand (Enriquez-Harris et al. 1991). This sequence weakly stimulates intrinsic pausing and enhances but is not required for 3' processing at the poly(A) site. Interestingly, it is functional in only one orientation. By analogy with TTF-1/Reb1p binding sites at the pol I terminator, this observation would suggest that the α2-globin pause site may be bound by a protein. However, no such terminator proteins have been found associated with this site nor is it predicted to promote secondary structure formation in RNA or DNA. The second example of a pause element is the palindrome sequence G_5AG_5 which is located between the closely spaced human compliment genes C2 and factor B (Ashfield et al. 1991; Yonaha and Proudfoot 1999). Originally, this site was characterized as a sequence that binds the zinc-finger transcription factor MAZ (Ashfield et al. 1994). It was first thought that MAZ acted as a roadblock to induce pausing, but it was later demonstrated that the G_5AG_5 sequence was able to weakly

induce intrinsic pausing and stimulate 3' processing in the absence of MAZ (Yonaha and Proudfoot 1999). The role of these two pause sites remains unclear. It is known that termination occurs within a few hundred bases of these sites unlike most pol II genes where termination occurs over thousands of nucleotides downstream of the poly(A) signal. This observation makes sense in the case of the C2 gene, where termination must occur close to the poly(A) site to prevent interference with the factor B promoter. Mechanistically, it could be that a weak pause site might give the polymerase some extra time to interact with the poly(A) site and 3' processing machinery. The 3' processing factors could then more effectively induce release (Zhao et al. 1999).

Yet another mRNA processing activity that is linked to pol II termination is splicing of the exon closest to the poly(A) site (Dye and Proudfoot 1999; Zhao et al. 1999). Specifically, the splice acceptor site that is located between the final intron/exon boundary is required for normal termination efficiency. Interestingly, it has been shown that the U2AF-65 splicing factor, which binds the poly-pyrimidine tract just upstream of the splice acceptor, also binds to the carboxy-terminus of the poly(A) polymerase (Vagner et al. 2000). Although once again the mechanistic details are not understood, it appears as if many of the mRNA processing events are initmately associated. It is known that factors involved in all steps of mRNA processing can bind to the CTD of pol II (McCracken et al. 1997; Zeng and Berget 2000; Fong and Bentley 2001). Thus, the CTD may functionally integrate signals from the processing machinery at all stages of transcription to influence the decision of the complex to terminate. Such a common link between 3' processing events may insure that termination does not occur prematurely and may prevent mRNAs that are not correctly processed from leaving the nucleus.

As previously mentioned, there are two categories of transcripts that are exceptions to the general scheme of pol II termination described above. First, are the transcripts encoding histone proteins, which are the only known mRNAs that are not polyadenylated. The 3' ends of histone transcripts are formed through a cleavage event that is signaled by a 16 nucleotide long stem-loop followed by a purine-rich binding site for the U7 snRNP. Cleavage occurs between the stem-loop and the U7 binding site while transcription continues for several hundred bases downstream before the polymerase terminates by a poorly defined mechanism. Although the details are unclear, it is known that termination in this case depends upon cleavage at the 3' processing site, which requires U7 snRNP and a stem-loop binding protein (SLBP) (Gick et al. 1986; Cotten et al. 1988; Vasserot et al. 1989; Chodchoy et al. 1991; Gu and Marzluff 1996). The second exception to the general termination mechanism is characterized by the U1, U2, U4, and U5 snRNAs (small nuclear RNAs). These non-polyadenylated RNAs are structural components of the snRNP particles that are involved in splicing pre-mRNA. The 3' ends of these particular snRNAs appear to be formed under the direction of a conserved signal sequence $GTTTN_{0-3}AAAPuNNAGA$ which is called the 3' box (Hernandez and Weiner 1986; de Vegvar et al. 1986). It was originally thought that this sequence may have directly stimulated termination. However, some new evidence suggests that this sequence directs endonucleolytic cleavage and trimming of the transcript while the polymerase continues transcribing downstream like it does in all other cases. Interestingly, the 3' box fails to direct termination if the transcript is not initiated from an authentic snRNA promoter. Limited footprinting analysis suggests that the region downstream of the U1 3' box is bound by an unidentified protein(s) that could

pause the polymerase (Cuello et al. 1999). Thus, snRNA termination is also not well understood. However, it seems that termination of histone and snRNA transcripts is dependent upon cleavage like other pol II transcripts.

In summary, the mechanism of pol II termination remains the least understood of all polymerases. However, it is apparent that nearly all events encompassing RNA synthesis are integrated through the CTD of pol II to influence termination efficiency (McCracken et al. 1997; McCracken et al. 1998). By requiring signals from the promoter and the 3' processing machinery, which must take into account alternative splicing and polyadenylation sites, the polymerase is prevented from releasing a transcript that is not correctly processed. In general, pol II termination is inhibited until the 3' processing machinery has begun to act upon signals located at the prospective 3' end. While the 3' end of the mature RNA is formed and modified, pol II continues to elongate the cleaved nascent transcript. Evidence suggests that the 3' processing machinery directly interacts with the polymerase to cause transcript release (McCracken et al. 1997; McCracken et al. 1998). However, the evidence does not rule out several other possibilities which include: 1) the loss of an anti-termination factor from the polymerase complex upon transcription of the processing signal, 2) a 5'-3' exonuclease that travels down the nascent transcript to dissociate the polymerase from the template, and 3) pausing induced by intrinsic sequences and/or DNA-binding proteins that enhance the action of other termination factors.

Elongation and Termination in Context

Knowing that there are many structural and functional similarities between all of the multi-subunit polymerases described above, the lack of conservation between the

termination mechanisms is initially somewhat surprising. However, the differences probably reflect overall strategies for polymerase regulation in the context of the entire organism. For example, bacteria need simple ways to rapidly coordinate the expression of several genes at once in response to metabolic requirements imposed by the environment. To accomplish this, many genes in the same metabolic pathway are expressed under one promoter as a polycistronic message (Lodish and Darnell 1995). As a supplementary control to the regulation of initiation, termination can be modulated by feedback inhibition loops which provide a straightforward way for bacteria to fine tune gene expression. As another example, pol III must efficiently recycle to accommodate the requirements of a growing cell for a large number of tRNAs and is thus highly proficient at termination. In contrast, pol II must be resistant to termination in order to synthesize very long, intricately modified transcripts. To prevent transcripts that have not been correctly processed from leaving the nucleus, pol II must be assured of transcript integrity by the processing machinery before termination can occur.

Vaccinia Introduction and Biology

Answering complicated questions about any process in molecular biology requires model systems that are genetically malleable. In the field of transcription, biologists have typically looked to bacteria, yeast, fruit flies, and some viruses as suitable models. The prototypic orthopoxvirus, vaccinia, lends itself particularly well to transcription studies because it quickly replicates in the cytoplasm of host cells and encodes nearly all of the factors required for nucleic acid metabolism on its 192kB double-stranded linear DNA genome (Conaway and Conaway 1994; Moss 1990a; Moss 1996). In addition, it provides unique and powerful genetic tools to specifically probe for

transcription factors, including those that are involved in termination. The background information in the remainder of this chapter is a prelude to the remainder of this work which is a characterization of the vaccinia J3 protein as a factor that influences 3' end formation.

There are three unique classes of vaccinia virus gene expression. Vaccinia genes are transcribed at early, intermediate, and late times post-infection by a virus encoded RNA polymerase that is homologous to the eukaryotic polymerases in terms of size and complexity (Baroudy and Moss 1980; Moss 1990b; Fields et al. 1996). Vaccinia transcription can be visualized as a cascade in which the products of each gene class initiate the expression of the succeeding class. Early gene transcription is promptly stimulated after viral entry by initiation factors contained within the infecting virion (Kates and Beeson 1970a; Kates and Beeson 1970b; Kates and McAuslan 1967). The products of early genes serve as intermediate transcription factors or as proteins involved in DNA replication. Intermediate gene expression begins after viral DNA replication has begun. Some of the known intermediate gene products are involved in transactivation of late gene expression. Late genes encode early gene transcription factors that are packaged into the virion for utilization during the next infection cycle (Moss 1996). Transcription initiation for each gene class is mediated by a unique upstream promoter element. Each promoter consists of core and initiator sequences that have been defined by saturation mutagenesis. All three types of promoters are A/T rich and are located in the region upstream of the transcription initiation site (Davison and Moss 1989a; Davison and Moss 1989b; Baldick et al. 1992). Like eukaryotic pol II transcripts, vaccinia mRNAs are capped (Wei and Moss 1975) and polyadenylated (Kates and

Beeson 1970a). However, in contrast to pol II transcripts, they are not spliced. Capping occurs during transcription when the nascent RNA reaches approximately 30 nucleotides in length. A viral capping enzyme removes the 5' terminal phosphate, catalyzes the 5' to 5' ligation of a GMP residue, and methylates the N^7 position of the terminal guanylate residue (Martin and Moss 1975; Martin and Moss 1976). Then, a separate bi-functional enzyme encoded by the J3 gene places a methyl group on the ribose moiety of the penultimate base (Barbosa and Moss 1978a; Barbosa and Moss 1978b). Finally, the transcript is polyadenylated by a virally encoded heterodimeric poly-A polymerase that contains the J3 2'-o-methyltransferase protein as a processivity subunit (Schnierle et al. , 1992).

Early transcription initiation requires the activity of two proteins encoded by the D6 and A7 genes (Gershon and Moss 1990). Together, they form a heterodimer called Vaccinia Early Transcription Factor, or VETF, which binds the early promoter core sequence (Broyles et al. 1991; Broyles and Li 1993). VETF has been found in complex with the RNA polymerase, which suggests that it may recruit the polymerase to the promoter (Li and Broyles 1993a). VETF has a DNA-dependent ATPase activity that is necessary for elongation (Li and Broyles 1993b). The polymerase, a 94-kDa RNA polymerase associated protein (RAP94, also known as H4), and VETF are required for reconstituting specific in vitro transcription initiation and elongation of early genes (Ahn et al. 1994; Gershon and Moss 1990). However, in vitro termination of early genes requires the activity of the mRNA capping enzyme, which is a heterodimer of the D1 and D12 proteins (Shuman et al. 1987; Broyles et al. 1988), and NPH-I (nucleoside triphosphate phosphohydrolase I; D11) which is a DNA-dependent ATPase (Mohamed

and Niles 2000). Termination occurs in response to a U_5NU signal located in the transcript. Upon transcribing the U_5NU termination sequence, the polymerase terminates approximately 30-50 nucleotides downstream (Yuen and Moss 1987). Although it is not known precisely how termination occurs, it is known that C-terminus of NPH-I binds to the H4 subunit of the polymerase and is required for transcript release (Mohamed and Niles 2000; Mohamed and Niles 2001). This mechanism of termination produces RNAs that are characteristically homogeneous in length.

Intermediate and late gene transcription is different from early transcription in a number of ways. First, intermediate and late gene expression cannot occur until viral DNA replication has begun (Vos and Stunnenberg 1988). Second, promoter sequences of known intermediate and late genes have been used to propose model core elements that are A/T rich and distinct both from early promoters and from one another (Baldick, Jr. et al. 1992). There is however, a TAAA initiator sequence that is shared by intermediate, late, and some early genes. This initiator sequence is hypothesized to cause the RNA polymerase to slip and catalyze the addition of a short poly-A leader sequence on the 5' end of transcripts (Ahn and Moss 1989; Ink and Pickup 1990). Third, intermediate and late genes have their own unique set of initiation factors. Most of these factors have been identified by reconstitution of in vitro transcription with successive fractionation of infected cell extracts. Intermediate factors include the capping enzyme (Vos et al. 1991) and two proteins called Vaccinia Intermediate Transcription Factors 1 and 2. VITF-1 was identified as a subunit of the RNA polymerase that is required in its free form, although it is normally found as a complex with the polymerase (Rosales et al. 1994). VITF-2 has been isolated from uninfected host nuclei, but its identity still remains a

mystery (Rosales et al. 1994). Several late transcription factors such as A1, A2, G8, and H5 have been discovered. A1, A2 and G8 were identified by a reverse genetic approach. Specifically, cloned vaccinia DNA fragments were co-transfected into infected cells with a β-galactosidase reporter gene driven by the 11K late promoter. The cells were treated with the DNA replication inhibitor araC to prevent expression of intermediate genes originating from the infecting virus. By combining the cloned fragments containing the A1, A2, and G8 genes, transcription from the reporter gene was reconstituted (Keck et al 1990). The H5 protein is an early gene product that has been purified from infected cell extracts and shown to increase the efficiency of in vitro late transcription (Kovacs and Moss 1996). Co-immunoprecipitation experiments suggest that H5 interacts with the transcription elongation factors G2 and A18 (Black et al. 1998). VLTF-X, a factor that was purified from uninfected cells, also acts to support late transcription in vitro (Gunasinghe et al. 1998; Wright et al. 1998).

One stark contrast between early versus intermediate and late transcription is the process of termination. A specific cis-acting termination sequence has not been identified for genes that are expressed after DNA replication has begun. The U_5NU terminator sequence that operates at the end of early genes can be found scattered throughout intermediate and late gene sequences, but apparently fails to be recognized by the postreplicativetranscription system. If a cis-acting terminator sequence exists for these genes, it must be degenerate in sequence and very inefficient in triggering 3' end formation. As a result, the polymerase reads through the 3' end of the intermediate or late coding sequence and terminates at numerous places downstream by an unknown mechanism. This termination system produces mRNAs that are heterogeneous in length

(Moss 1996). The 3' ends of the transcripts can be found anywhere from one to four kilo-bases downstream of the promoter. Compared to early genes, much less is known about the mechanism by which intermediate and late gene transcription termination occurs. However, genetic experiments have implicated three proteins in the process of postreplicative transcription elongation and termination. These proteins are the products of the A18, G2, and J3 genes (Bayliss and Condit 1993; Bayliss and Condit 1995; Black and Condit 1996; Lackner and Condit 2000; Latner et al. 2000; Xiang et al. 2000a).

A18

In a previous search for genes involved in controlling the pattern of gene expression, sixty-five temperature sensitive mutants were isolated. These mutants were placed into thirty-two complementation groups and were screened for defects in DNA and protein synthesis. All of the mutants in one complementation group had an "abortive late phenotype" that was characterized at the non-permissive temperature by protein synthesis that begins normally then stops abruptly at a late time point. Two of these abortive late mutants, *ts22* and *ts23*, were chosen for further characterization (Condit and Motyczka 1981; Condit et al. 1983; Pacha and Condit 1985).

Marker rescue experiments mapped the C*ts22* and *Cts23* mutations to the A18 open reading frame (Pacha et al. 1990). The protein product is expected to be 57-kDa and has some homology to the helicase motifs of the ERCC3 subunit of the pol II transcription factor TFIIH. Northern blots demonstrated that the A18 gene is expressed early and late during infection. Later work showed that the A18 protein is present in the core particles and has DNA dependent ATPase and 3'-5' DNA helicase activities (Bayliss and Condit 1995; Simpson and Condit 1994; Simpson and Condit 1995). Bayliss et al

discovered that an A18 mutant infection is characterized by promiscuous or aberrant transcription (Bayliss and Condit 1993). Promiscuous transcription is defined as transcription from regions of the genome that are not usually transcribed at late times in a wild type infection. Xiang et al demonstrated that the observed promiscuous transcription was due to readthrough from an upstream intermediate or late gene rather than early promoter reinitiation or random initiation (Xiang et al. 1998). These readthrough transcripts can hybridize to transcripts originating from downstream promoters that initiate transcription in the opposite direction. This phenomenon creates double stranded RNA, which can induce the host 2-5A pathway of RNA degradation catalyzed by RNase L (Bayliss and Condit 1993). Degradation of the transcripts results in the abortion of late protein synthesis (Pacha and Condit 1985). Recently, Lackner and Condit demonstrated that the A18 protein is required in conjunction with an unknown cellular factor to permit transcript release from the template in an in vitro transcription assay (Lackner and Condit 2000). These observations indicate that A18 acts as a postreplicative negative transcription elongation factor or more specifically, as a transcript release factor.

IBT

Isatin-β-thiosemicarbazone (IBT) is an anti-poxviral drug that is structurally composed of a six member and a five member ring with a short sulfur-containing side chain and is thus similar to a purine (Katz 1987). Although the neither the target nor the mechanism of IBT action is known, treatment of a wild type infection with IBT results in a readthrough transcription phenotype, similar to that observed during an A18 mutant infection (Meis and Condit 1991). Due to a high spontaneous mutation rate, approximately one in 10^{-4}-10^{-5} viruses in a wildtype infection will be either resistant to or

dependent upon IBT for growth. Several mutations conferring IBT dependence have been mapped to transcription factors and are discussed below. For technical reasons, the IBT resistance mutations are more difficult to map, but it is known that resistance can result from mutation of the 132-kDa subunit of the polymerase (Condit et al. 1991), the J3 2'-o-methyltransferase, and the G2 transcription elongation factor. IBT can therefore be used as a tool to identify proteins that are involved in transcription.

G2

It was originally hypothesized that the direct target of IBT may be the A18 protein because mutation of A18 results in a readthrough transcription phenotype that is similar to treatment of a wildtype infection with IBT. Thus, several IBT dependent viruses were identified and mapped by marker rescue with the expectation that A18 mutations would be discovered. Surprisingly, none of the IBT dependent (IBT^d) mutations mapped to A18. Each IBT^d mutation mapped to the G2 gene instead (Meis and Condit 1991). The 26-kDa G2 protein is expressed early during infection, has no previously characterized function, and is not homologous to any known proteins. It was shown that the G2 mutants have normal early protein expression, but synthesize reduced amounts of large proteins at late times post-infection. It was postulated that the late large protein synthesis defect was the result of aberrant transcription because the G2 mutants are dependent upon IBT for growth. Analysis of steady state mRNA isolated from cells infected with a G2 mutant in the absence of IBT revealed that the mutant virus produces shorter than wild type transcripts that are specifically truncated from the 3' end (Black and Condit 1996). This observation logically follows the IBT dependent phenotype. IBT presumably

compensates for the short RNA synthesis defect in the G2 mutants and promotes synthesis of transcripts that are of normal length.

Given that the mRNA synthesis phenotype of the G2 IBT^d mutants is so distinctly opposite from that of the A18 mutants, it was proposed that a G2 null mutation could serve as an extragenic suppressor of the *Cts23* mutant A18 allele. Viable A18-G2 double mutant viruses were isolated in two different ways. First, double mutants were isolated from a mixed infection with both *Cts23* and G2 deletion mutant viruses. Second, viable A18-G2 double-mutants were isolated from a screen of *Cts23* phenotypic revertants. The viability of the A18-G2 double mutant viruses suggest that each mutation compensates for the defect in the other gene product (Condit et al. 1996). In summary, these observations indicate that G2 is a positive postreplicative transcription elongation factor. An interesting twist in the search for A18-G2 double mutants was revealed when one of the *Cts23* phenotypic revertants was shown to contain a mutation that did not map to G2 as predicted. The characterization of this suppressing allele has brought about the characterization of the J3 protein as a positive transcription elongation factor that is similar to G2 and is described in the following chapters.

CHAPTER 2
CHARACTERIZATION OF J3 AS A TRANSCRIPTION FACTOR

Introduction

As described in the previous chapter, the observation that distinctly opposite phenotypes are produced from defects in the transcription factors A18 and G2 led to the hypothesis that mutations in both genes could suppress each other. As mentioned, viable A18-G2 double mutants have been obtained by two different methods (Condit et al. 1996). First, a G2 deletion mutant was recombined with the *Cts23* mutant allele in a mixed infection. Second, mutant G2 alleles were found as second-site suppressors of *Cts23*. During the search for G2 mutations as suppressors of *Cts23*, a virus called *r51* was isolated that grew at 40°C, was IBT sensitive, and formed slightly smaller than wildtype plaques (Latner et al. 2000). Previous experience with other G2 suppressor mutations indicated that the suppressing allele in *r51* might generate IBT dependence. Therefore, *r51* was crossed to wildtype virus to segregate the suppressor mutation from the A18 mutation. An IBT dependent virus called *r51x4* was isolated from this cross and its mutation was mapped by marker rescue. Surprisingly, the IBT^{d} mutation in *r51x4* did not map to the G2 gene. Instead, it was shown that the J3 gene in *r51x4* was responsible for IBT dependence. Sequence analysis revealed that the J3 gene in *r51x4* had, in fact, two mutations. The upstream mutation resulted in the exchange of a glycine at codon 96 with an aspartic acid residue (G96D). The downstream mutation resulted in the exchange of an arginine for a lysine (R327K) in the portion of J3 that overlaps the adjacent J4

open-reading frame. Marker rescue with sub-clones of the J3 gene demonstrated that the upstream G96D mutation was responsible for generating the IBT dependent phenotype. A virus containing only the G96D mutation was isolated by crossing *r51x4* to wildtype virus and by screening the IBT dependent progeny by PCR and restriction fragment length polymorphism analysis. This single mutant virus was designated *J3x* and is phenotypically identical to *r51x4*. To prove that the G96D mutation was responsible for the suppression of the *Cts23* A18 allele, the *J3x* virus was crossed with the *Cts23* virus. A viable double-mutant virus called *J3x23* was isolated and is phenotypically the same as *r51*. Like wildtype virus, *r51* and *J3x23* are ts^+ and IBT sensitive. However, both mutants form smaller than wildtype plaques and are difficult to grow to high titer. Given that the *J3x* mutation can cause IBT dependence and suppress *Cts23* just like mutations in G2, it was hypothesized that J3 may also be a positive postreplicative transcription elongation factor. Moreover, the realization that IBT dependence can arise from mutation of another gene besides G2 suggested that other unidentified transcription factors could be discovered by selecting for additional IBT dependent viruses (Latner et al. 2000; Xiang et al. 2000a). This chapter reports the isolation and mapping of several IBT dependent viruses and the characterization of J3 as a positive transcription elongation factor.

Materials and Methods

Cells and Viruses

The continuous African green monkey kidney cell line BSC40 and conditions for cell culture, vaccinia virus cultivation, infection, plaque assay, and one-step growth have been previously described (Condit and Motyczka 1981; Condit et al. 1983). Wildtype

vaccinia virus strain WR, the G2 gene mutant *G2A*, the J3 mutants *J3x*, *J3-7*, and *r51x4*, the genes J3 and A18 double-mutant *J3x23*, and the conditions for their growth, infection, and plaque titration have been previously described (Condit and Motyczka 1981; Condit et al. 1983; Meis and Condit 1991; Black and Condit 1996; Latner et al. 2000; Xiang et al. 2000a). Briefly, *G2A* is phenotypically IBT dependent (IBT^d) and contains an engineered 10-bp deletion in gene G2, resulting in a frameshift at codon 90 and truncation of the 220-amino acid protein at position 93. *J3x* is phenotypically IBT^d and contains a missense mutation in codon 96 of the J3 gene (G96D). *J3-7* is phenotypically IBT^d and contains a single nucleotide deletion in the J3 gene that causes a frameshift at codon 49 and truncation of the 333-amino acid protein at position 58. *r51x4* is phenotypically IBT^d and contains two missense mutations in the J3 gene, one in codon 96 of gene J3 (G96D), and the other in codon 327 of J3 (R327K) within an overlap with the N terminus J4 open reading frame, where it affects codon 22 of gene J4 (D22N). *J3x23* is phenotypically wt [ts^+ and IBT sensitive (IBT^s)], but forms smaller plaques than wt. *J3x23* is a recombinant between *J3x* and the gene A18 ts mutant *Cts23*. IBT was prepared fresh before each use and applied at a final concentration of 45μM as previously described (Pacha and Condit 1985).

Protein Pulse-Labeling

Pulse-labeling of proteins in virus infected cells was done as described previously (Condit and Motyczka 1981). Briefly, confluent 35-mm dishes of BSC40 cells were infected with *wt* or mutant virus at a high multiplicity of infection (m.o.i.) in the absence of IBT and incubated at 37 or 40°C as indicated in the figure legends. At various times post-infection, cells were metabolically labeled with [^{35}S]methionine for 15 min. Cells

were lysed on the dishes by addition of SDS-polyacrylamide gel electrophoresis (PAGE) sample buffer, and solubilized proteins were analyzed by SDS-PAGE. The gels were Coomassie blue stained, dried, and autoradiographed.

Isolation of Total Cellular RNA

Isolation of total cellular RNA was performed as previously described (Xiang et al. 2000b). Briefly however, confluent 100 mm dishes of BSC40 cells were infected with wt or mutant virus at a high m.o.i. in the absence of IBT, and incubated at 37°C or 40°C as indicated in figure legends. At various times after infection, total cellular RNA was purified using RNeasy Total RNA purification columns as described by the manufacturer (Qiagen, Inc., Chatsworth, CA). The RNA was eluted from column with DEPC-treated H2O and quantified by measuring absorbance at 260 nm.

Northern Analysis

Northern analysis was performed as previously described (Xiang et al. 2000b). Northern transfers were prepared as follows. Two μg of purified RNA in water was adjusted to a final concentration of 0.36x MOPS buffer (1x MOPS buffer = 20 mM MOPS, 8 mM sodium acetate, 1 mM EDTA, pH 7.0), 1.6 M formaldehyde, 36 % formamide, in a final volume of 13.75 μl. The samples were heated to 65°C for 15 minutes, then chilled on ice. One μl of a 1mg/ml ethidium bromide and 2 μl of RNA loading buffer (50 % glycerol, 1 mM EDTA, 0.0 5% bromophenol blue, 0.0 5% xylene cyanol) were added to each sample prior to loading on the gel. The samples were electrophoresed at 80 volts through 1.2 % agarose gels containing a final concentration of 1 x MOPS buffer and 0.37 M formaldehyde in an electrophoresis buffer containing 1 x MOPS buffer and 0.23 M formaldehyde. RNA was transferred to GeneScreen membrane

(New England Nuclear) in a buffer containing 30 mM sodium phosphate, pH 6.5, and the RNA was UV crosslinked to the membrane. Membranes were hybridized with either uniformly labeled riboprobes, or with 5' end labeled DNA oligonucleotides as described below.

Riboprobes were synthesized essentially as described in the Promega Protocols and Applications Guide. Reactions (35 µl) contained 1x optimized transcription buffer (Promega), 11 mM DTT, 40 units RNase inhibitor (Promega), 0.4 mM ATP, GTP, and UTP, 11 µM unlabeled CTP, 50 µCi [α^{32}P]-CTP (3000Ci/mmol), 1 µg linearized template DNA, and 20 units of T7 RNA polymerase. Reactions were incubated at 37°C for one hour. One unit of RNase-free DNase (Promega) was added and the reactions were incubated for 15 minutes at 37°C. 65 µl of STE/SDS (10mM Tris-HCl pH 8.0, 1mM EDTA, 0.1M NaCl, 0.1%SDS) was added and the samples were extracted with an equal volume of phenol:chloroform. The aqueous phase was precipitated by addition of 100 µl 4 M ammonium acetate, 120 µl isopropanol, and 1 µl of 20 mg/ml glycogen, and incubation at RT for 30 min. The precipitate was collected by centrifugation and resuspended in 50 µl DEPC treated water.

Anti-sense DNA oligonucleotide probes (45-56mers) used for northern analysis of F17 gene transcripts were labeled as follows: lyophilized oligonucleotides (Genemed Synthesis Inc., San Francisco, CA) were resuspended in water at a final concentration of 200 ng/µl. 200 ng of oligonucleotide was added to 11 µl water and warmed to 70°C for one minute, then chilled on ice. Reactions were adjusted to a final volume of 20 µl by addition of 2 µl of 10x T_4 polynucleotide kinase buffer (Promega), 50 µCi [γ-^{32}P]-ATP (6000Ci/mmol), and 10 units of T_4 polynucleotide kinase (Promega). Reactions were

incubated at 37°C for 1 h. 1 µl of 0.5 M EDTA was added to stop the reaction followed by the addition of 125 µl of TE (10 mM Tris-HCl, pH 7.5, 1 mM EDTA). Unincorporated label was removed by two centrifugations through sephadex G25 spun columns essentially as described and the total TCA precipitable radioactivity in the final eluate was determined.

For hybridization with riboprobes, membranes were pre-hybridized for 2 hours at 55°C in 10 ml of buffer containing 50 mM Tris-HCl pH 7.5, 1 M NaCl, 50 % formamide, 1 % SDS, 0.1 % sodium pyrophosphate, 10x Denhardt's [0.2 % BSA, 0.2 % polyvinylpyrolidone, 0.2 % Ficoll], 10 % dextran sulfate, 0.1 mg/ml salmon sperm DNA. Membranes were then hybridized overnight at 55°C in the same buffer containing 10^7 cpm of riboprobe. Following hybridization, membranes were washed once at room temperature in 0.1 % SDS, 0.1x SSC, three times at 65°C in 1.0 % SDS, 0.1x SSC, and exposed to film.

For hybridization with labeled DNA oligonucleotides, membranes were prehybridized in 10 ml of buffer containing 6x SSC, 0.1 % SDS, 10x Denhardt's reagent, and 100 ug/ml salmon sperm DNA for 2 hours at 42°C, then hybridized overnight in the same buffer containing 10^7cpm of labeled oligonucleotide probe. Membranes were washed twice for 20 minutes at 50°C in 5x SSC, 0.1 % SDS and exposed to film. The amount of label hybridized to each sample was quantified after autoradiography using a phosphorimager (Molecular Dynamics).

DNA Sequence Analysis

Sequence analysis was performed as described (Latner et al. 2000). Briefly sequence of mutant virus DNA was obtained by sequencing PCR products amplified

from genomic DNA which was isolated by one of two different methods: (1) DNA was isolated as described except the amounts were scaled down twenty-fold (Esposito et al. 1981). (2) Qiagen DNeasy miniprep spin columns (Qiagen Inc., Santa Clarita, CA) were utilized to purify total infected cell DNA from 200 ul of infected cell lysate following the manufacturer's instructions for isolating DNA from cells in culture. The complete J3 and G2 coding sequences were PCR amplified with two different sets of primers that hybridize just outside of the open reading frames. The J3 primer pair yields a 1197-bp product and the G2 primer pair yields a 1067-bp product. Sequence was obtained from both strands of DNA with the same primers used for amplification and with two additional primers that hybridize in the middle of the coding sequence to give overlapping products. Sequencing was performed by the University of Florida Interdisciplinary Center for Biotechnology Research (ICBR) DNA Sequencing Core Laboratory.

RFLP Analysis

RFLP analysis was performed as described (Latner et al. 2000). Briefly, the screen for the *J3x* single mutant virus from the *r51x4* testcross was performed by restriction fragment length polymorphism analysis of two different PCR products generated from virus genomic DNA. Candidate mutant plaques were picked, grown, and viral genomic DNA prepared as described above. A 379 bp PCR product that spanned the G96D upstream mutation was amplified with appropriate primers, then digested with Eag I. A downstream 522 bp PCR product that spanned the mutation in the J3/J4 overlapping region (J3:R327K/J4:D22N) was amplified with appropriate primers, then digested with Bst YI. Restriction fragments were analyzed on 1 % MetaPhor agarose (FMC BioProducts, Rockland, ME) in TBE. Both restriction enzymes cleave wild-type

sequence but not mutant sequence. A mutant virus was identified that was resistant to Eag I in the 379 bp upstream PCR product but sensitive to Bst YI in the downstream 522 bp PCR product. Subsequent sequence analysis of the entire J3 open-reading frame confirmed the presence of the upstream mutation and absence of the downstream mutation. The single mutant virus was designated *J3x* and was utilized for further characterization.

DNA Clones & Marker Rescue

Marker rescue was performed as previously described (Meis and Condit 1991; Thompson and Condit 1986; Latner et al. 2000). The cosmid clones used for initial marker rescue experiments have been previously characterized (Thompson and Condit 1986). The J fragment subclones pJ5, pJ6, and pJ7 used for J3 marker rescue were generously supplied by Jerry Weir and have been described (Ensinger et al. 1985). The clone pJ3R, which contains the precise J3 open reading frame cloned in pET14b, was a gift from Ed Niles (SUNY Buffalo).

Western Blot Analysis

Western analysis was performed as previously described (Latner et al. 2000). 125 ul of virus lysate was combined with 250 ul of SDS-PAGE sample buffer. 35 ul of each sample was electrophoresed on an 8 % polyacrylamide-SDS gel then the proteins were transferred from the gel to nitrocellulose in 25 mM Tris-Base, 192 mM glycine, 20 % methanol for 3 hours at 4°C and 80 volts. The membrane was blocked with 5 % dry milk in PBS for 1.5 hours, washed 3 times for 5 min with 0.05 % NP-40 in PBS, and incubated for 1 hour with anti A18, G2 and J3 primary antibodies diluted 1:10,000, 1:1,000, and 1:5,000 respectively in PBS plus 5 % dry milk. The monoclonal anti-A18 and G2 antibodies have been previously described (Black and Condit 1996). The rabbit

polyclonal anti-J3 antibody was a gift from Ed Niles (SUNY Buffalo). The blot was then washed 3 times for 5 min with 0.05 % NP-40 in PBS and incubated with secondary HRP conjugated anti-rabbit and anti-mouse antibodies diluted 1:10,000 and 1:5,000 respectively for 1 hour. The blot was washed 3 times for 10 min with PBS plus 0.05 % NP-40 and once for 2 min in PBS. Final detection was performed with an ECL western detection kit (Amersham Pharmacia) according to the manufacturer's instructions.

Results

Isolation of IBT-Dependent Mutants

Given that previously characterized mutants of the G2 gene are dependent upon IBT for growth, additional genes encoding transcription elongation factors were sought by selecting for novel spontaneous IBT-dependent mutants. Wildtype virus was plaqued at 37°C in the absence of drug and 10 well-isolated plaques were picked and grown to create working stocks. This step was performed to ensure that any mutants derived from these wt stocks would not be related to one another. The 10 wt lysates were then titered in the presence and in the absence of IBT at 37°C. In each of the 10 wt stocks, several spontaneous mutant viruses that formed plaques in the presence of IBT were observed. For each of the 10 original wt stocks, 10 plaques that grew in the presence of drug were picked, giving a total of 100 viruses capable of growing in the presence of IBT. Next, these 100 viruses were plaqued in the presence and in the absence of drug to determine whether they were resistant to or dependent upon IBT for growth. At least one IBT-dependent virus isolate was identified from 9 of the 10 original wt virus stocks. One IBT-dependent plaque arising from each of these 9 wt stocks was chosen for in-depth study. The mutation in each IBT-dependent virus was mapped by marker rescue in a

fashion similar to *J3x* and the region containing the mutation was sequenced. The results, summarized in tables 2-1 and 2-2, show that two of the mutants contain deletions in gene G2 and seven of the mutants contain alterations in gene J3. *G2-5* contains a 3-nucleotide in-frame deletion that removes a valine from position 105 of the G2 gene. *G2-2* contains a single base deletion that results in a frameshift at codon 209 and a premature termination, yielding a truncated polypeptide 213 amino acids in length. All of the IBT-dependent J3 mutants contain insertions, deletions, or nonsense mutations that result in truncation of the protein significantly from its wt length of 333 amino acids. Interestingly, each of the J3 insertion and deletion mutations is located in a region of the coding sequence that contains a long stretch of A or T residues (table 2-3). Presumably, these hot spots for deletion and insertion arise from slipped misparing during DNA replication. One of the J3 mutant viruses, *J3-7*, was chosen for further characterization. *J3-7* contains a single base deletion at codon 49, resulting in a frameshift and a truncation of the protein to 58 amino acids in length. In summary, these results confirm that, like G2, null mutation of gene J3 results in IBT dependence and by analogy with G2 indicates that the J3 gene plays a role in the regulation of postreplicative transcription elongation during vaccinia virus infection.

Western Blot Analysis of A18, J3, and G2 Proteins in Mutant Infections

Western blot analysis was performed to confirm predictions derived from sequence data regarding expression of the A18, G2, and J3 proteins in each mutant virus infection (figure 2-1). Infected cell lysates were electrophoresed through 8% SDS polyacrylamide gels and transferred to nitrocellulose. The blots were simultaneously probed with anti-A18, anti-G2, and anti-J3 antibodies and developed by chemiluminescence. All three proteins are readily apparent in wildtype virus lysates.

Cts23 makes all three proteins, but the relative amount of A18 protein is reduced, consistent with previous results (Simpson and Condit 1994). The *G2A* control virus encodes a frameshift in the G2 gene and, as previously described (Black et al. 1998), fails to produce G2 protein. The *J3x* missense mutant produces detectable J3 protein while the other IBT-dependent J3 truncation mutants, *J3-1*, *J3-3*, *J3-4*, *J3-7*, *J3-8*, *J3-9*, and *J3-10*, produce no observable J3 protein. No G2 was observed in *G2-2*, which is missing a valine from position 105, or in *G2-5*, which has a frameshift truncation of the C-terminal 11 amino acids of the G2 protein, indicating that these mutant proteins are unstable. A18, J3, and G2 proteins are all detectable in the *J3x23* double-mutant but at very low levels, presumably because this particular virus grows poorly and is therefore present in the lysate at a very low concentration. In addition, the relative amount of A18 protein in the *J3x23* lysate is low, consistent with the presence of the *Cts23* mutation in this virus. In summary, the results of the Western blot analysis are consistent with the DNA sequence analysis of these viruses, showing that the J3 protein of the *J3x* mutant is stable and that the new IBT^d J3 and G2 mutants produce no detectable J3 or G2 protein.

Plaque Assay

Plaque phenotypes of the *J3x*, *J3x23*, and *J3-7* viruses are compared to parental viruses in the assay shown in figure 2-2. Wildtype virus forms large plaques at 37°C and intermediate-sized plaques at 31 and 40°C and is IBT sensitive. *Cts23* is both temperature sensitive and IBT sensitive. *r51*, the phenotypic revertant of *Cts23* and parent of *J3x*, is ts^+ and IBT sensitive, but forms smaller than wt plaques at both 31 and 40°C. *J3x* is IBT dependent and phenotypically identical to its parent *r51x4* (data not shown), confirming that the downstream J3 R327K mutation in *r514* does not contribute to the *J3x* mutant

plaque phenotype. *J3x* forms larger plaques at 37°C relative to 31 and 40°C, but is somewhat leaky with respect to IBT dependence at both 31 and 37°C. *J3x23* is phenotypically identical to *r51*, confirming that the J3 G96D mutation in *J3x* is sufficient to suppress *Cts23* and that the downstream J3 R327K mutation in *r51* does not contribute to the plaque phenotype of *r51*. *J3-7*, which contains a frameshifting deletion of the J3 gene, has a noticeably different phenotype when compared to the point mutant *J3x* virus. *J3-7* forms smaller plaques than *J3x* at 37°C and is tighter with respect to IBT dependence at 37°C. Since *J3-7* is likely to be a null mutant, this observation suggests that the *J3x* virus may retain some J3 activity that contributes to its growth phenotype.

One-Step Growth

To characterize the growth phenotypes of each virus in more detail, one-step growth experiments were performed. BSC40 cells were infected at an m.o.i. of 10 with wt, *Cts23*, or *J3-7*, or at an m.o.i. of 5 with *J3x23*. The infections were incubated under permissive and nonpermissive conditions appropriate for each virus, harvested at various times postinfection, and then titered under permissive conditions appropriate for each virus (figure 2-3). Wildtype virus growth was compared to *J3x* growth at 40°C in the presence and in the absence of IBT (figure 2-3 panel A). At 40°C, wildtype virus yielded 1-5 PFU/cell, significantly higher than the T_0 background, and growth was inhibited by IBT. *J3x* produced only background levels of progeny in the absence of IBT and approximately 1 PFU/cell in the presence of drug, slightly higher than background. This result confirms and extends the plaque assay, showing that *J3x* is IBT dependent and difficult to grow to high titer. *J3-7* growth was compared to wildtype at 37°C in the presence and in the absence of IBT (figure 2-3 panel B). Wildtype grew to very high titer

in the absence of drug, and growth was inhibited by IBT. *J3-7* yielded less than 1 PFU/cell in the absence of drug and slightly greater than 1 PFU/cell in the presence of IBT, indicating that under one-step growth conditions *J3-7* is moderately IBT dependent and confirming that, like *J3x*, it is also difficult to grow. The growth of *J3x23* was compared to wt and *Cts23* at both 31 and 40°C in the absence of IBT. Both wt and *Cts23* produced 4-10 PFU/cell at 31°C, but *Cts23* produced only background levels of virus at 40°C, consistent with plaque assay results. *J3x23* generated 0.5-1 PFU/cell at both temperatures, which is higher than background, but less than wildtype. These results confirm that the *J3x* mutation suppresses *Cts23* and show that *J3x23* grows very poorly. In summary, these experiments show that the one-step growth properties of all three mutant viruses are generally consistent with the plaque assay phenotype and that the mutant viruses all grow poorly under permissive conditions.

Viral Protein Synthesis

As an initial test of the prediction that the J3 mutant viruses would have an *in vivo* biochemical phenotype similar to that of the previously described G2 mutants, a time course of gene expression from the *J3x*, *J3-7*, *and J3x23* viruses was compared with that of the wt virus by metabolic labeling of proteins in infected cells. Infected cells were pulse labeled with [^{35}S]methionine at various times postinfection; the total infected cell protein was analyzed by SDS-PAGE and autoradiography (figure 2-4). In the wt infection, host protein synthesis is shut off by 4 hours post-infection (p.i.), synthesis of early proteins begins by 2 h p.i. and subsequently decays, and synthesis of intermediate and late viral proteins begins by 4 h p.i. and persists for the duration of the experiment. In the *J3x* infection, host shutoff and synthesis of early viral proteins are

indistinguishable from those in the wt infection. Intermediate and late *J3x* protein synthesis is initiated at the appropriate time relative to a wt infection, and low-MW proteins are synthesized in normal amounts; however, large late proteins are synthesized in reduced amounts relative to small *J3x* proteins and relative to the wt infection. In the *J3-7* infection, the schedule of host shutoff and early and late proteins synthesis is somewhat delayed compared to the schedule in the wt infection. Early *J3-7* proteins and low MW *J3-7* intermediate and late proteins are synthesized in slightly reduced quantities; however, large late *J3-7* proteins are synthesized in greatly reduced amounts relative to small late *J3-7* proteins and relative to a wt infection. In the *J3x23* infection, the schedule of host shutoff and viral gene expression is delayed compared to that of either the wt or the J3 single-mutant infections, however, proteins are synthesized in relatively normal amounts. The delayed schedule of protein synthesis observed in the *J3x23* infection likely results from the fact that this infection was done at relatively low m.o.i. and thus may be somewhat asynchronous. Most important, large late *J3x23* proteins, although reduced somewhat in quantity relative to a wt infection, are synthesized in increased amounts relative to the J3 single mutant infections. In summary, the most prominent effect of J3 gene mutation on viral protein synthesis is a reduction in synthesis of large late proteins, identical to the phenotype previously reported for G2 mutants (Black and Condit 1996). The increase in synthesis of large late proteins observed in *J3x23* shows that suppression of the *J3x* mutation by the A18 *Cts23* mutation correlates with restoration of a wt protein synthesis phenotype.

Viral mRNA Synthesis

It was hypothesized that the defect in protein synthesis observed in J3 mutant infections is caused by synthesis of shorter than normal intermediate and late RNAs as

had been previously demonstrated for G2 mutants. This hypothesis was addressed by conducting Northern blot analysis of mutant RNAs. Cells were infected with wt or mutant viruses, and total infected cell RNA was prepared at various times postinfection, electrophoresed on denaturing formaldehyde agarose gels, transferred to nylon membranes, and hybridized with standard early (gene C11), intermediate (gene G8), and late (gene F17) gene riboprobes (figure 2-5). As previously described, wt virus produces early RNAs that are homogeneous in size due to discrete transcription initiation and termination events, whereas intermediate and late RNAs are heterogeneous in size due to variation in 3' end sequence. Expression of the early RNA in a wt infection begins immediately after infection, peaks at ~3 h, and subsequently decreases to undetectable levels. Wildtype intermediate and late RNAs appear between 3 and 6 h postinfection and persist throughout the experiment. In a *J3x* infection, the pattern of early RNA synthesis is indistinguishable from the wt infection in size, quantity, and kinetics. *J3x* intermediate and late RNAs are heterogeneous in size, and they are synthesized in similar amounts and with the same kinetics as a wt infection; however, the average chain length of both the intermediate and late *J3x* RNAs is decreased relative to a wt infection. In the *J3-7* infection, the schedule of synthesis of all three classes of RNA is slightly delayed relative to a wt infection, but the amounts of RNA synthesized are similar to those in a wt infection. The *J3-7* early RNA is identical in size to wt RNA. Most important, like the *J3x* infection, the *J3-7* intermediate and late RNAs, although still heterogeneous in size, have an average chain length that is decreased compared to wt RNA and similar to *J3x* RNA. In the *J3x23* infection, schedule of synthesis of all three classes of RNA is delayed significantly relative to the wt infection, and the RNAs are synthesized in decreased

amounts, consistent with the *J3x23* pattern of protein synthesis and diagnostic of a lower m.o.i., slightly asynchronous infection. The *J3x23* early RNA is the same size as wt RNA. Importantly, the intermediate and late *J3x23* RNAs, although somewhat reduced in chain length relative to wt RNA, are nevertheless increased in size relative to the J3 single-mutant infections. In summary, mutation of the J3 gene results in synthesis of intermediate and late RNAs that are reduced in size relative to a wt infection, identical to what is observed in a G2 mutant infection (Black and Condit 1996). Suppression of the *J3x* mutation by the A18 *Cts23* mutation correlates with partial restoration of a wt RNA synthesis phenotype.

Analysis of Late Viral mRNA 3' Ends

Based on the comparison with G2 mutants, it was hypothesized that J3 mutants would produce intermediate or late transcripts that are truncated specifically from their 3' ends. To test this hypothesis, wt and mutant F17 gene transcripts isolated 12 h p.i. were once again inspected by Northern analysis, this time using as probes labeled anti-sense DNA oligonucleotides that hybridize to various places downstream of the F17 promoter (figure 2-6). The G2 null mutant virus *G2A* was included in this experiment as an additional control. It is noteworthy that the wt F17 late RNA is somewhat unusual compared to most other late RNAs in that a relatively strong homogeneous 1.4-kb transcript can be detected above the more heterogeneous background. This transcript is detected with probes a-e and not with probe f, indicating that its 3' end is localized within the region between 1160 and 1510 downstream from the F17 promoter. It has been determined that this predominant homogeneous 3' end is due to transcript cleavage by an unknown virus induced factor (Susan D'Costa personal communication). Cleavage of poxvirus transcripts is quite rare, but has been described for the cowpox ATI transcript

(Howard et al. 1999). Regardless of the cleavage event, F17 transcripts can be used to gauge the average length of transcripts produced by each of the mutant viruses as follows: Hybridization of probe a to wt and mutant RNAs shows that the F17 transcripts are reduced in size in the *J3x*, *J3-7*, and *G2A* infections but normal in size in the *J3x23* infections, consistent with the Northern blots described above. Close inspection of the hybridization with probe a reveals discrete bands within the population of *J3x*, *J3-7*, and *G2A* RNAs and also that the *J3x* RNAs are generally larger than the *J3-7* and *G2A* RNAs. As hybridization probes are moved progressively further downstream from the F17 promoter, hybridization to the smaller J3 and G2 mutant transcripts is lost until, with probe e, only the longest transcripts remain, corresponding to the full-length major F17 transcript. This pattern of hybridization indicates that all of the short mutant transcripts originate from the F17 promoter; that is, the short F17 mutant transcripts are truncated from their 3' ends. Quantitative analysis of F17 transcription (figure 2-6 panel C) shows that the total amount of hybridization to *J3x*, *J3-7*, or *G2A* RNA is equal to or greater than wt near the F17 promoter and decreases relative to the wt signal as a function of the distance from the F17 promoter. Thus the quantification shows that initiation of transcription of the F17 gene in the J3 mutant infections occurs at normal levels and confirms that the F17 RNAs are 3' truncated. By contrast, the amount of hybridization to *J3x23* RNA, although reduced relative to wt RNA, remains constant as a function of distance from the F17 promoter, consistent with rescue of the J3 mutant phenotype in the A18-J3 double mutant virus. Together, these observations show that the J3 single mutants produce 3' truncated late transcripts and together with other evidence presented support the hypothesis that the J3 protein affects the elongation or termination of

intermediate and late gene transcription. The existence of discrete RNA species (bands) within the 3' truncated RNAs from mutant infections could reflect preferred promoter proximal transcription pause or termination sites revealed by mutation of an elongation factor. The fact that *J3x* RNAs are slightly longer than *J3-7* or *G2A* RNAs may indicate that relative to the null mutants, the *J3x* point mutant retains some elongation factor activity.

Discussion

Based on previous phenotypic analysis of mutants in the vaccinia virus G2 postreplicative positive transcription elongation factor, two independent genetic selections have been used to search for additional vaccinia genes involved in the regulation of viral transcription elongation (Latner et al. 2000). In one selection, the *J3x* G96D mutation was isolated as an extragenic suppressor of the *Cts23* mutation in the A18 transcript release factor. In the second selection, nine viruses were isolated that contain spontaneous mutations which render the viruses dependent upon the anti-poxviral drug IBT for growth. Consistent with previous results, two of these mutants have defects in the G2 elongation factor. Mutations in the other seven IBT^d selected viruses mapped to the J3 gene. Examination of these J3 mutants reveals that although they have normal early gene transcription, they produce short, 3' truncated postreplicative transcripts. The production of 3' truncated postreplicative transcripts by the J3 mutants is consistent with the absence of a positive transcription elongation factor (or anti-termination factor) in the mutant infections. Generalized 3' truncation of postreplicative transcripts would account for the observed synthesis of abnormally short intermediate and late RNAs, and translation of abnormally short mRNAs would account for the observed specific decrease

in synthesis of large but not small intermediate and late proteins. The phenotype is also consistent with genetic data: growth of the J3 mutants can be rescued by procedures that enhance transcription elongation, specifically either treatment with the anti-poxviral drug IBT or recombination with mutants in gene A18. By analogy with previous characterization of the G2 gene (Meis and Condit 1991; Black and Condit 1996; Condit et al. 1996), these results strongly suggest that the J3 gene is an essential positive regulator of postreplicative gene transcription elongation.

The two different selections described, IBT dependence and *Cts23* suppression, may discriminate the two different types of J3 mutants isolated. Specifically, selection for IBT dependence resulted in isolation entirely of null J3 mutants, while selection of *Cts23* resulted in isolation of the more subtle *J3x* missense mutation, which, as described below, may retain some J3 protein activity. It is noteworthy in this regard that while mutation of J3 suppresses *Cts23* as judged by plaque formation, the resulting double-mutant virus grows extremely poorly, as judged by one-step growth. Since J3 null mutants are more defective in growth than *J3x*, it seems likely that null mutation of J3, while it may compensate for the effects of *Cts23*, would leave a double-mutant virus too crippled to grow at all, thus prohibiting isolation of J3 null mutant suppressors of *Cts23*. Consistent with this idea, several attempts at constructing a *J3-7-Cts23* double-mutant by recombination have been unsuccessful. By contrast with the J3 gene, previous results show that null mutants of the G2 gene can suppress temperature-sensitive mutants in A18, including both *Cts23* and *Cts22* (Condit et al. 1996). In addition, Western blot analysis of three previously isolated G2 point mutant suppressors of *Cts23* and *Cts22* (*r41*, *cs1*, and *cs4*) has shown that none of these mutants produce detectable G2 protein

(unpublished results). These observations could suggest that the G2 gene has a more limited range of function than the J3 gene, such that a knockout of the G2 function can be tolerated in the presence of an A18 gene mutation.

Interestingly the J3 protein, also known as VP39 (vaccinia protein-39kDa), has two previously characterized activities that modify both the 5' and 3' end of viral transcripts. First, J3 is a (nucleoside-2'-o-)methytransferase that uses S-adenosylmethionine (SAM) to methylate the 2' position of the 5' penultimate ribose found in cap 0 mRNAs, thereby generating a cap 1 structure (illustrated in figure 2-7) (Barbosa and Moss 1978a; Barbosa and Moss 1978b; Shi et al. 1996). Although the cap-0 structure has been shown to be important for mRNA stability and for efficient initiation of translation, the purpose of the 2'-o-methylated cap 1 structure has yet to be determined. J3 protein can be found in infected cells as a monomer and as a heterodimer complexed with the virus encoded E1 poly(A) polymerase (Schnierle et al. 1992). Both the monomer and the heterodimer have (nucleoside-2'-o-)methytransferase activity. The second function of J3 is to provide a stimulatory activity to the E1 poly(A) polymerase (Schnierle et al. 1992). *In vitro* experiments have shown that E1 catalyzes the polyadenylation of RNA 3' ends in two distinct and separable phases (Gershon 1998). In the absence of J3 protein, E1 rapidly adds 30-35 A residues in an abrupt burst of polymerization, then it converts to a much slower, non-processive mode of poly-A synthesis. Upon addition of the J3 subunit, the polymerase converts once more to a highly processive form that is competent for rapid poly(A) addition. Together, E1 and J3 rapidly synthesize poly(A) tails in vivo that are up to 200 nucleotides long. Interestingly, the J3 crystal structure has been published and has revealed that the J3 molecule is an

oblate sphere with a characteristic cleft along one side (Hodel et al. 1996a). Previous work has shown that the (nucleoside-2'-o-)methyltransferase activity resides on one surface of the molecule and the binding site for E1 is found on the exact opposite face of the molecule (figure 2-8) (Shi et al. 1997). Knowing the structural relationships between the methyltransferase and E1 binding sites, it is no surprise that published site-directed mutagenesis studies have shown that the (nucleoside-2'-o-)methyltransferase and E1 stimulatory activities can be genetically separated (Gershon et al. 1998). Thus, J3 is known to modify both the 5' and 3' ends of mRNAs by providing at least two distinct activities that are functionally independent. The new observation that J3 positively affects postreplicative gene transcription elongation, as described above, is not inconsistent with the intimate association between 3' processing and termination of pol II transcription and begs the following question: Can the J3 transcription elongation activity be separated from the other two mRNA modification activities of the protein or are there any functional links?

The nature of the J3 mutations described above provides some initial clues to the relationship between the putative J3 elongation factor activity and the other two previously described activities. Both DNA sequence and Western blot analysis of the seven new J3 mutants selected initially for IBT dependence show that these are null mutants: each contains a chain-terminating nonsense or frameshift mutation, and none synthesizes detectable J3 protein. Thus all of these seven mutants should be lacking in both poly(A) polymerase stimulatory and (nucleoside-2'-o-)methyltransferase activities. By contrast, *J3x*, which was initially isolated as a *Cts23* suppressor but which is also IBT dependent, contains a missense mutation (G96D) that does not affect the steady state

levels of J3 protein produced during infection. Under permissive conditions *J3x* forms larger plaques than a representative J3 null mutation (*J3-7*), suggesting that the *J3x* mutant may retain some J3 protein function. Interestingly, the published crystal structure of J3 shows that the *J3x* G96D mutation is localized very near the methyltransferase active site, between a highly conserved aspartic acid at position 95 and an arginine at position 97, both of which form hydrogen bonds with the methyl donor S-adenosylmethionine (Hodel et al. 1996b). Replacement of the glycine, which forms a kink in the alpha carbon backbone of J3 between the two charged residues, with an additional aspartic acid residue could easily compromise binding of SAM to the enzyme. Therefore, if the *J3x* protein is missing either of the previously identified J3 activities, it seems most likely that it would be defective in methyltransferase and retain poly(A) polymerase stimulatory activity.

Published preliminary experiments (Xiang et al. 2000a), conducted by Dr. Ying Xiang and not shown here, have demonstrated that purified recombinant *J3x* protein is in fact defective for methyltransferase activity *in vitro*. This result raised the intriguing possibility that methyltransferase activity is somehow functionally coupled to transcription elongation or termination. However, it did not prove a link between the methyltransferase and transcription because of two remaining possibilities. First, it could be that a cellular methyltransferase could replace the J3 activity *in vivo*, thus making it impossible to link the two functions. Second, it could be that in addition to disrupting methyltransferase activity, the G96D mutation could affect the surface charge of the protein, and thus coincidentally affect interactions with other proteins hypothetically involved in regulation of transcription elongation or termination. In addition to

examining the methyltransferase activity of *J3x*, it was also shown that the *J3x* protein is normal for both *in vitro* and *in vivo* E1 stimulation and makes normal length poly(A) tails. In contrast, it was shown that the *J3-7* null mutant, as hypothesized, is defective for E1 stimulation as it makes short poly(A) tails *in vivo* (Xiang et al. 2000a). This result preliminarily indicated that E1 stimulation is not linked to the transcription elongation activity since both *J3x* and *J3-7* both make short 3' truncated transcripts. To more directly investigate the relationship between the methyltransferase and transcription elongation activities and to prove that E1 stimulation is not linked to J3 transcription factor activity, it became necessary to construct site-directed mutants in attempt to genetically and functionally segregate the three J3 activities as described in the next chapter.

All of the analyses described in this chapter that were used to examine J3 mutant transcripts measure steady-state RNA *in vivo* and therefore do not formally distinguish whether the differences in RNA structure observed by comparing wt and J3 mutant RNAs is attributable to abnormal RNA synthesis or to abnormal RNA turnover. In both yeast and mammalian cells, RNA turnover mechanisms exist which degrade RNA exonucleolytically from either the 5' or 3' end (Gallie 1998; van Hoof and Parker 1999). Interestingly, the 5' to 3' pathway is influenced by both the RNA cap structure and the poly(A) binding protein and thus bears provocative similarities to the system described here that involves a viral function that interacts with both RNA 5' and 3' ends. However, the 3' truncation of RNA observed here is both specific and limited, leaving wt quantities of promoter proximal RNA in J3-mutant-infected cells. The persistence of normal quantities of relatively stable 3'-truncated RNA seems inconsistent with any of the

degradation pathways heretofore described, thus the interpretation that the J3 protein affects synthesis rather than turnover of RNA is favored.

Infections with J3 mutant viruses resulted in early gene expression that is indistinguishable from that in a wt virus infection, demonstrating that the influence of the J3 gene on transcription elongation is highly specific for intermediate and late genes. This observation emphasizes further the fact that early poxvirus transcription elongation and termination are mechanistically distinct from intermediate and late transcription elongation and termination, which in turn seem to be mechanistically similar. The vaccinia RNA polymerase exists in two forms, one that transcribes early but not late genes and one that transcribes late genes but not early genes. The late-gene-specific RNA polymerase contains eight core virus-coded subunits only (Wright and Coroneos 1995). The enzyme that transcribes early genes contains an additional subunit, the H4 (Rap94) gene product, which is required for initiation at early promoters and which remains tightly associated with the RNA polymerase during elongation (Ahn et al. 1994; Deng and Shuman 1994). Thus the subunit composition of early and late elongating RNA polymerases reflects a memory of the class of gene that is being transcribed. Early genes terminate transcription efficiently at discrete sites in response to a specific *cis*-acing signal in an energy-dependent process that involves the viral heterodimeric capping enzyme (genes D1 ad D12) and a viral DNA-dependent ATPase (gene D11) (Shuman et al. 1987; Christen et al. 1998; Deng and Shuman 1998). The postreplicative elongation machinery ignores early termination signals and terminates inefficiently at numerous sites of unknown sequence composition (Cooper et al. 1981; Mahr and Roberts 1984). Genetic and biochemical analyses suggest that elongation and termination of

postreplicative genes are influenced by as many as five different factors, four of which are virus coded, one of which is host coded, and all of which are distinct from factors regulating early gene transcription elongation and termination. Postreplicative elongation is negatively regulated *in vivo* by the A18 gene product, which contains DNA-dependent ATPase (Bayliss and Condit 1995), DNA helicase (Simpson and Condit 1995), and transcript release activities (Lackner and Condit 2000), and thus behaves like a transcription termination factor. A18-catalyzed transcript release *in vitro* requires an additional factor found in uninfected cells (Lackner and Condit 2000), thus implicating a host factor in postreplicative vaccinia transcription termination. Positive transcription elongation factors now include both the J3 and G2 gene products (Black and Condit 1996; Latner et al. 2000; Xiang et al. 2000a), the latter of which has no other known function. Whether J3 and G2 contribute essential activities to the same process or are functionally redundant remains to be determined. The G2 protein associates specifically with the H5 gene product (Black et al. 1998), an abundant phosphoprotein implicated in both late transcription and morphogenesis (Kovacs and Moss 1996; Beaud and Beaud 1997; DeMasi and Traktman 2000). Although the J3 protein does not specifically associate with viral RNA polymerase isolated from virions (Broyles and Moss 1987), evidence suggests that A18, G2, and H5 may associate with a larger complex *in vivo* during infection (Black et al. 1998), raising the possibility that many or all of the postreplicative viral transcription elongation/termination factors may exist in a large elongation complex. A wide variety of vertebrate poxviruses, including variola (Massung et al. 1994), alastrim (Shchelkunov et al. 2000), molluscum contagiosium (Senkevich et al. 1997), Shope fibroma (Willer et al. 1999), and myxoma (Cameron et al.

1999), contain all three essential postreplicative elongation factors, A18, J3, and G2, implying that the genes coevolved from a primordial vertebrate poxvirus as an integrated elongation regulatory complex. Interestingly, the insect poxviruses *Melanoplus sanguinipes* Entomopoxvirus (Afonso et al. 1999) and *Amsacta moorei* Entomopoxvirus (Bawden et al. 2000) each contain both A18 and J3 but lack G2, suggesting that late gene transcription may have evolved slightly differently for the insect poxviruses.

Despite the specificity of the J3 elongation activity for postreplicative genes, the multifunctional J3 protein is present throughout infection (Nevins and Joklik 1977), packaged in virions (Moss et al. 1975), and thus provides cap ribose methylase and poly(A) stimulatory activities during the early as well as postreplicative stages of vaccinia infection. Interestingly, the A18 (Simpson and Condit 1994) and H5 (Paula Traktman, personal communication) proteins are also present in virions, though their implied role in early gene expression remains to be determined. It is conceivable therefore that the early and postreplicative transcription elongation complexes, while displaying differences in composition and activity, also overlap in structure and function. By contrast, the G2 gene product is synthesized early during infection, is present late during infection, but is not packaged in virions and thus seems to be an exclusively postreplicative gene specific factor (Meis and Condit 1991; Black et al. 1998). Further comparisons of the structure and activity of the early and postreplication transcription elongation complexes should reveal interesting similarities required for proper elongation of all viral genes, and differences that account for the apparently different termination mechanisms.

Table 2-1: G2 IBT[d] Mutants

Isloate	Mutation [a]	Protein [b]
Wild type	wt	wt; 220aa
G2-5	deletion (3)	V105Δ; 219aa
G2-2	deletion (1)	*fs* 209; 213aa

[a] The number of nucleotides affected by each mutation are included in parenthesis
[b] "*fs*-# " indicates the codon within which a frame shift occurs. "Δ" indicates deletion.

Table 2-2: J3 IBT[d] Mutants

Isolate	Mutation [a]	Protein [b]
Wild type	wt	wt; 333aa
J3-1	insertion (1)	*fs* 50; 50aa
J3-3	insertion (1)	*fs* 50; 50aa
J3-4	nonsense	W183*; 182aa
J3-7	deletion (1)	*fs* 49; 58aa
J3-8	insertion (1)	*fs* 34; 50aa
J3-9	nonsense	S5*; 4aa
J3-10	deletion (1)	*fs* 235; 246aa

[a] The number of nucleotides affected by each mutation are included in parenthesis
[b] "*fs*-# " indicates the codon within which a frame shift occurs. "*" indicates a stop codon.

Table 2-3: Sequence context of J3 deletion and insertion mutations

Isolate		Sequence Context
wt	135	AGAATTATTTTTT.CTTAGTAA
J3-1	135	AGAATTATTTTTTTCTTAGTAA
J3-3	135	AGAATTATTTTTTTCTTAGTAA
J3-7	135	AGAATTATTTTT..CTTAGTAA
wt	87	GGTCGCAAAAAAA.CTGCCGTA
J3-8	87	GGTCGCAAAAAAAACTGCCGTA
wt	689	TAAATTATGAAAAAAAGATGTA
J3-10	689	TAAATTATGAAAAAA.GATGTA

Sequence of J3 insertion and deletion mutants from three different regions of the J3 gene are compared in each case to wt sequence. The number to the left of the sequence is the number of the first nucleotide in the sequence, setting the A in the J3 ATG initiation codon as +1. Dots have been inserted into either wt or mutant sequence to facilitate alignment as necessary.

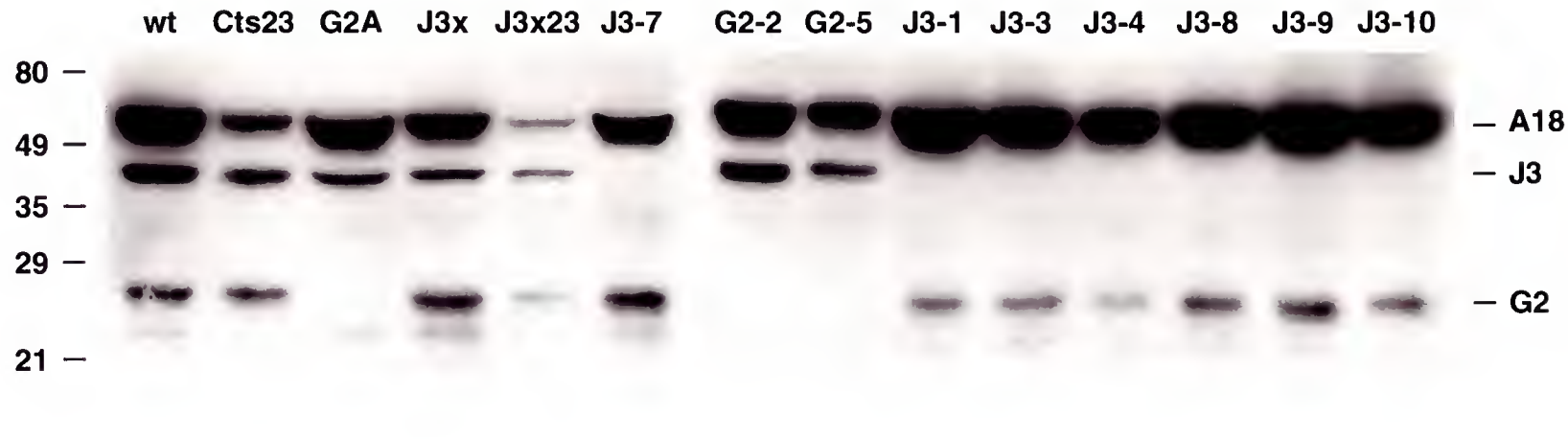


Figure 2-1. Western blot analysis of A18, J3, and G2 proteins synthesized during wt and mutant infections. Infected cell lysates were electrophoresed on SDS-PAGE, proteins were transferred to membranes, and membranes were probed simultaneously with anti-A18, anti-J3, and anti-G2 sera and developed by chemiluminescence as described under Materials and Methods. The migration of molecular mass standards, in kDa, is shown the left. The identity of the mutant virus used for infection is shown at the top. The migration of the specific protein detected is shown at the right.

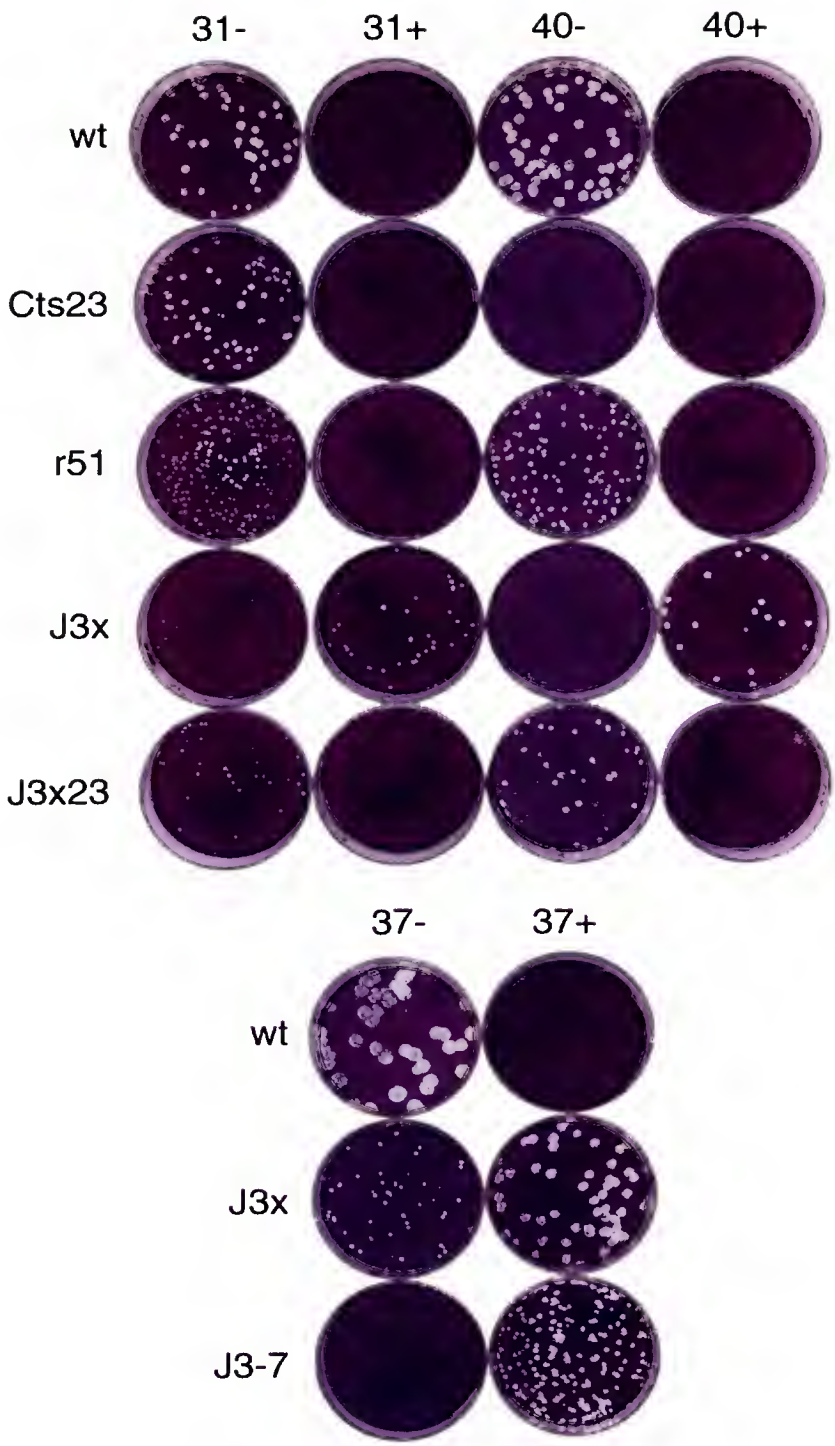


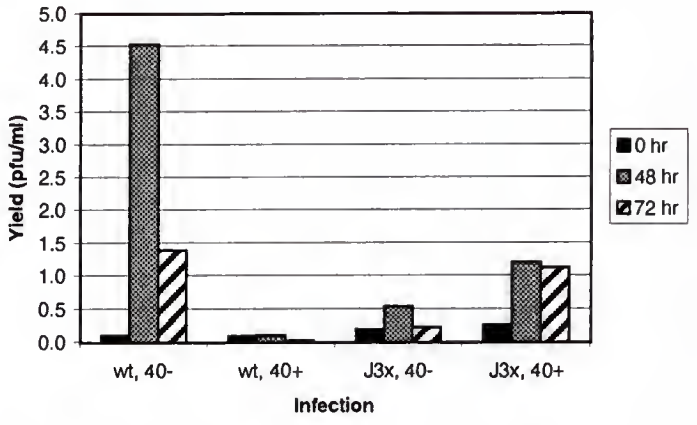
Figure 2-2. Plaque phenotypes of mutant viruses. Confluent monolayers of BSC40 cells in 60-mm dishes were infected with an appropriate dilution of virus and incubated in the presence or in the absence of IBT under an agar overlay for 6 days. Dishes were stained overnight with a second neutral red containing agar overlay. Agar was then removed and cells were stained with crystal violet. The mutant used for infection is indicated at the left of each row. Each column is labeled at the top to show the temperature of incubation (31, 37, or 40°C) and whether or not IBT was included (+ or -).

Figure 2-3. One-step growth analysis of J3 mutant viruses. Confluent monolayers of BSC40 cells in 35-mm dishes were infected with the viruses indicated at a m.o.i. of 10 for wt, *Cts23*, *J3x*, and *J3-7* or at an m.o.i. of 5 for *J3x23*. Infections were incubated in liquid medium under varying conditions of temperature (31, 37, or 40°C) in the presence or in the absence of IBT (+ or -) and harvested at 0, 48, and 72 h, as indicated. Lysates were harvested and plaque titrated under permissive conditions appropriate for each virus: wt, 37°C –IBT; *Cts23*, 31°C –IBT; *J3x*, 40°C +IBT; *J3-7*, 37°C +IBT; *J3x23*, 40°C –IBT. In each panel, numerical data are presented at the left, and bar graphs of the same data are shown on the right. (A) *J3x* is compared to wt. (B) *J3-7* is compared to wt. Note in the bar graph in B that the wt data have been truncated at 2 PFU/cell to aid in visualization of the mutant data. (C) *J3x23* is compared to wt and *Cts23*.

A

J3x

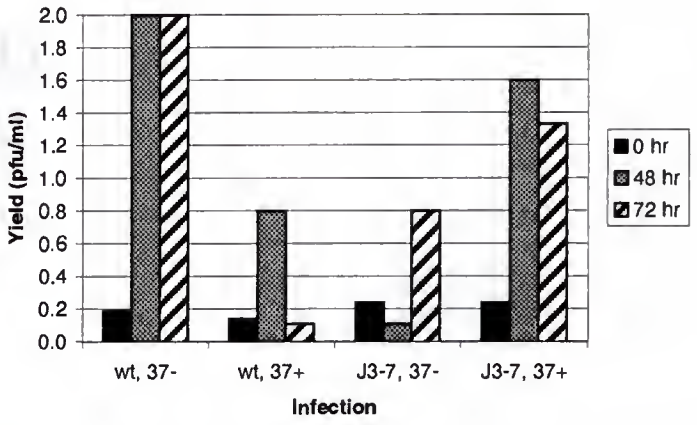
Infection	Yield (pfu/cell) 0 hr	48 hr	72 hr
wt, 40-	0.10	4.53	1.39
wt, 40+	0.09	0.10	0.02
J3x, 40-	0.19	0.53	0.22
J3x, 40+	0.26	1.20	1.12



B

J3-7

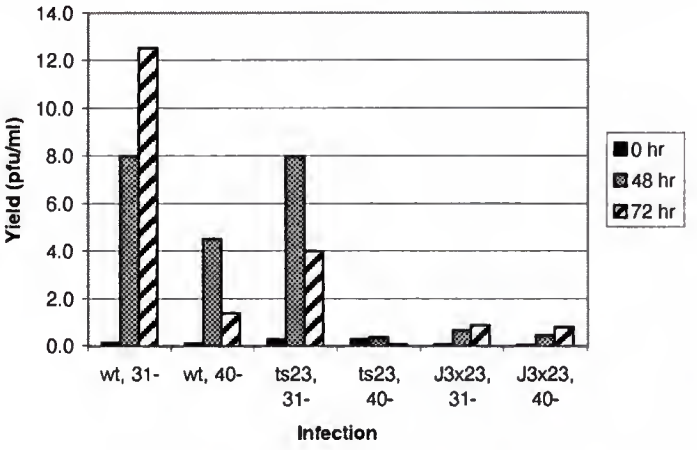
Infection	Yield (pfu/cell) 0 hr	48 hr	72 hr
wt, 37-	0.19	184.00	186.67
wt, 37+	0.14	0.80	0.11
J3-7, 37-	0.24	0.11	0.80
J3-7, 37+	0.24	1.60	1.33



C

J3x23

Infection	Yield (pfu/cell) 0 hr	48 hr	72 hr
wt, 31-	0.15	8.00	12.53
wt, 40-	0.10	4.53	1.39
ts23, 31-	0.27	8.00	4.00
ts23, 40-	0.27	0.35	0.05
J3x23, 31-	0.05	0.67	0.88
J3x23, 40-	0.04	0.43	0.80



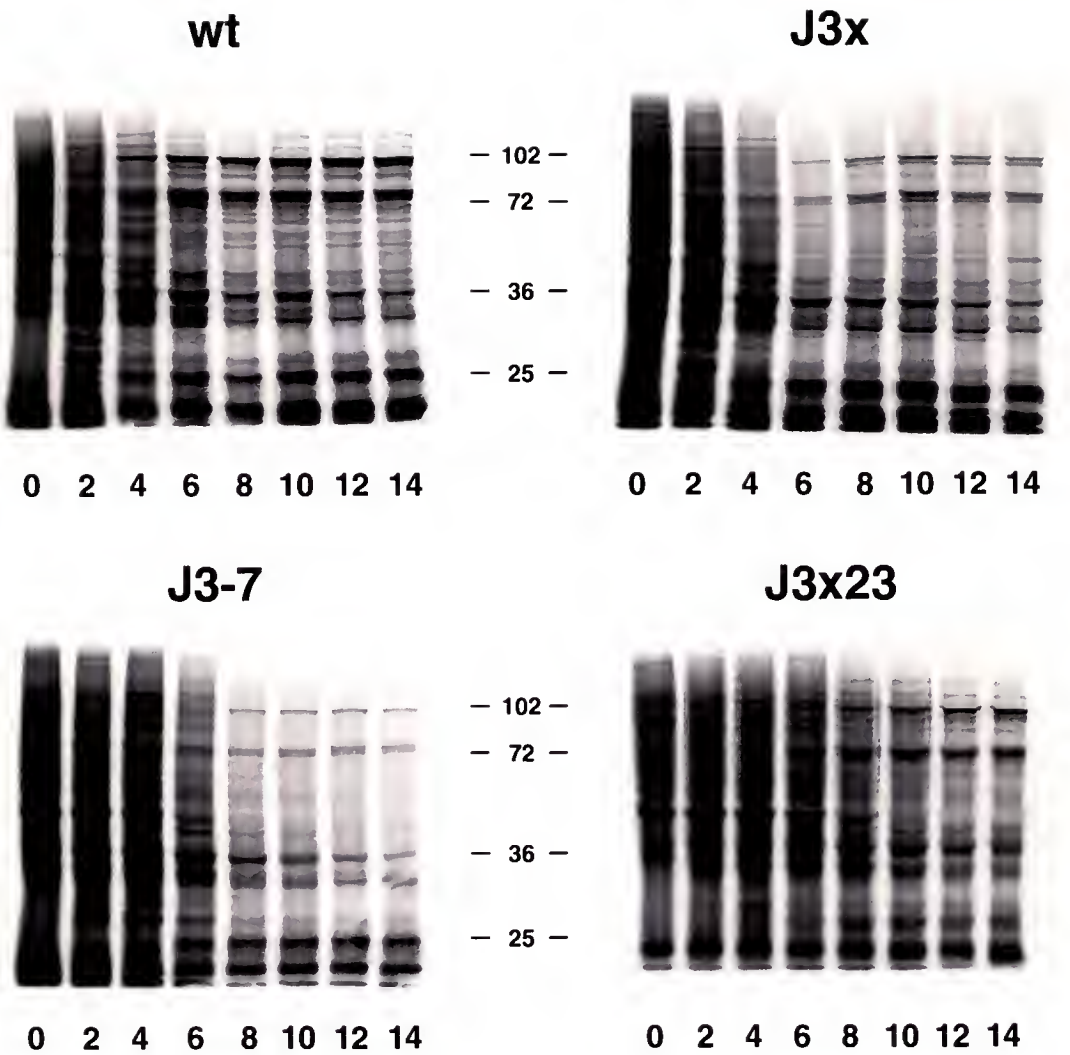


Figure 2-4. Protein synthesis in wt and mutant-infected cells. Confluent monolayers of BSC40 cells in 35-mm dishes were infected at m.o.i.=10 with wt, J3x, or J3-7, or m.o.i.=5 with J3x23, incubated at 40°C (wt, J3x, J3x23) or 37°C (J3-7), and pulse labeled for 15 min with [35S]methionine at various times after infection. Pulse-labeled proteins were analyzed by SDS-PAGE and autoradiography; autoradiographs are shown. The mutant used for the infection is indicated at the top of each autoradiogram. The time at which the pulse label was done is indicated, in hours, at the bottom of each autoradiogram. The migration of molecular weight markers, in kDa, is indicated between the autoradiograms.

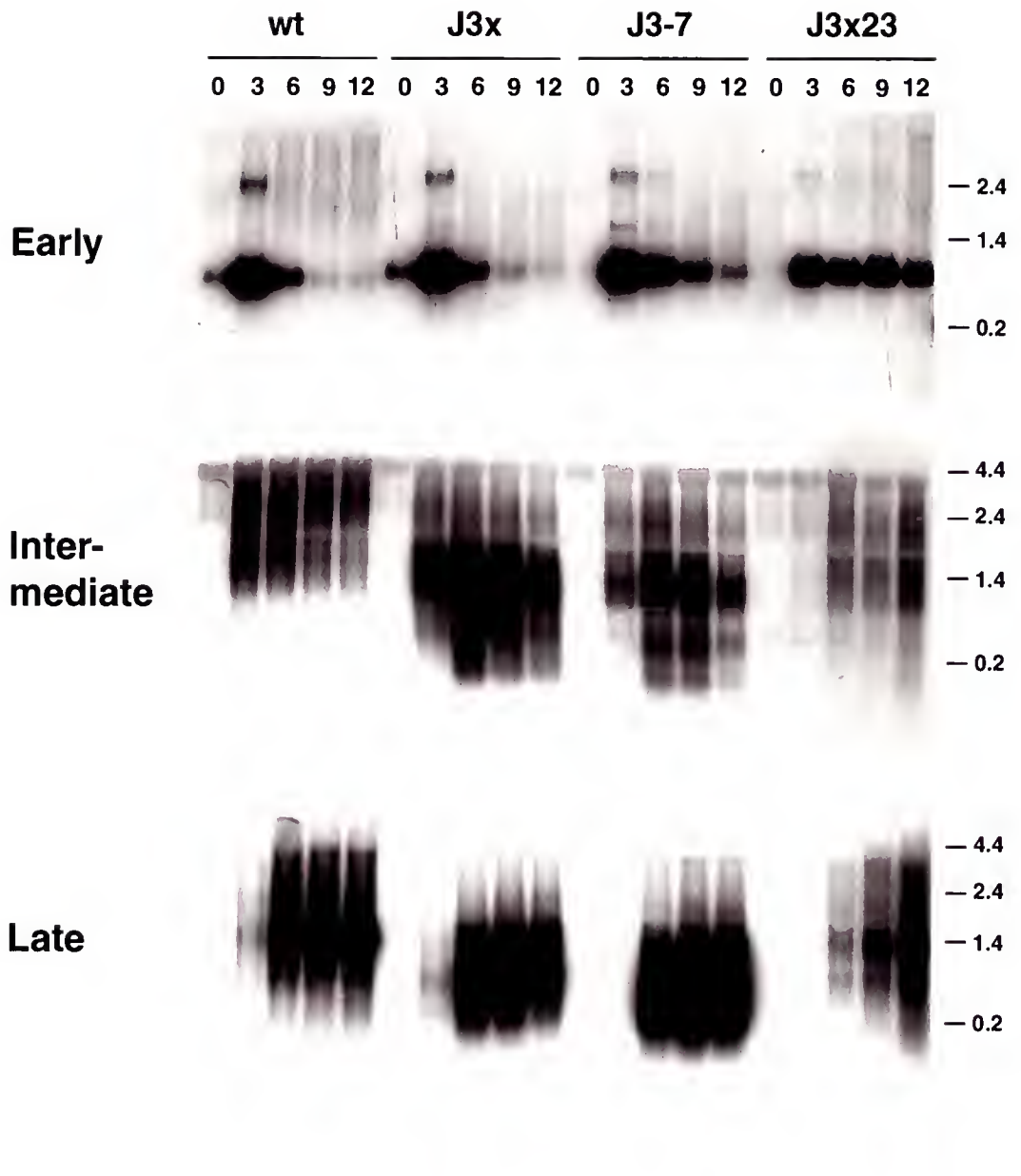
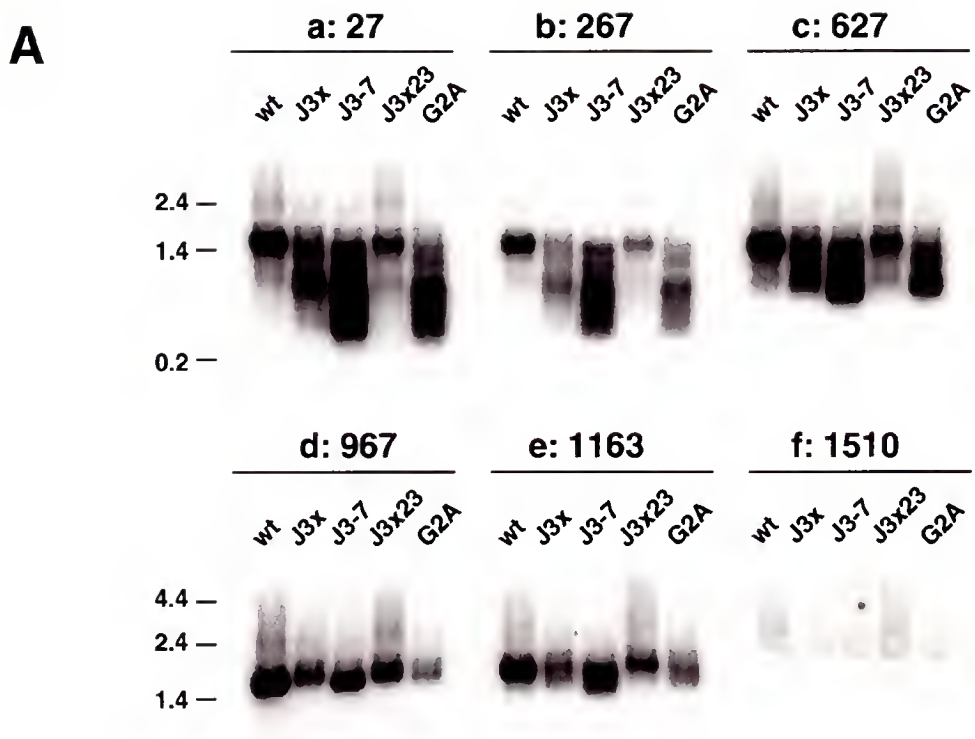
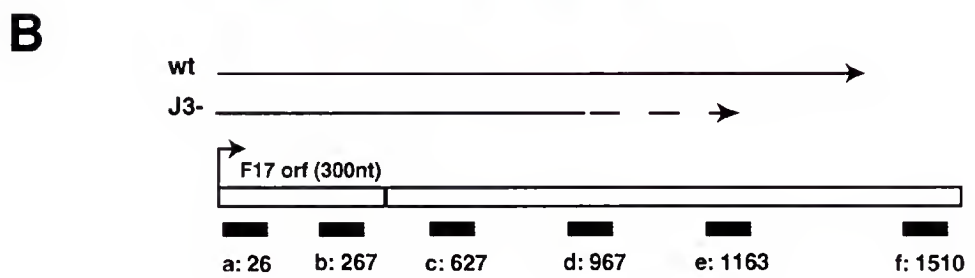


Figure 2-5. Northern analysis of RNA from wt and mutant-infected cells. Confluent monolayers of BSC40 cells in 100-mm dishes were infected at m.o.i. =10 with wt, J3x, or J3-7, or m.o.i.=5 with J3x23, incubated at 40°C (wt, J3x, J3x23) or 37°C (J3-7). Total cellular RNA was purified from infected cells at various times postinfection, indicated in hours above the autoradiograms, transferred to nylon membranes and hybridized with radiolabeled antisense riboprobes specific for an early (gene C11), intermediate (gene G8), or late (gene F17) gene, as indicated to the left of the autoradiograms. The migration of size markers, in kb, is indicated to the right of the autoradiograms.

Figure 2-6. Structural analysis of the gene F17 late RNA from wt and mutant-infected cells. (A) Confluent monolayers of BSC40 cells in 100-mm dishes were infected at m.o.i.=10 with wt, *J3x*, *J3-7*, or *G2A*, or m.o.i.=5 with *J3x23*, incubated at 40°C (wt, *J3x*, *J3x23*, and *G2A*) or 37°C (*J3-7*) as indicated above the autoradiograms. Total cellular RNA was purified from infected cells at 12 h p.i., transferred to nylon membranes, and hybridized with radiolabeled antisense DNA oligonucleotides a-f, as indicated above the autoradiograms. The position of the middle of each oligonucleotide (average length 50 nt) relative to the F17 promoter is shown. [In the F17 gene, the first transcribed nucleotide (+1) is coincident with the A in the translation initiation ATG.] The migration of size markers, in kb, is indicated to the left of the autoradiograms. (B): A cartoon showing (from top to bottom) an interpretation of the results shown in (A), with wt and J3 mutant (J3-) transcripts indicated as arrows (representation of the J3 mutant transcripts as a broken arrow indicates heterogeneity at the 3' end); a schematic of the F17 transcription unit with the promoter indicated as an arrow, the 300-nu coding sequence indicated as the leftmost open box, and the remainder of the sequence indicated as the rightmost open box; the map positions of oligonucleotide probes a-f. (C) Quantification of the data in (A). For each probe, the total radioactivity in each lane was measured, and each mutant lane was then plotted as a percentage of the wt signal for a given probe.


A
a: 27
b: 267
c: 627
wt J3x J3-7 J3x23 G2A
2.4
1.4
0.2
d: 967
e: 1163
f: 1510
4.4
2.4
1.4


B
wt
J3-
F17 orf (300nt)
a: 26
b: 267
c: 627
d: 967
e: 1163
f: 1510

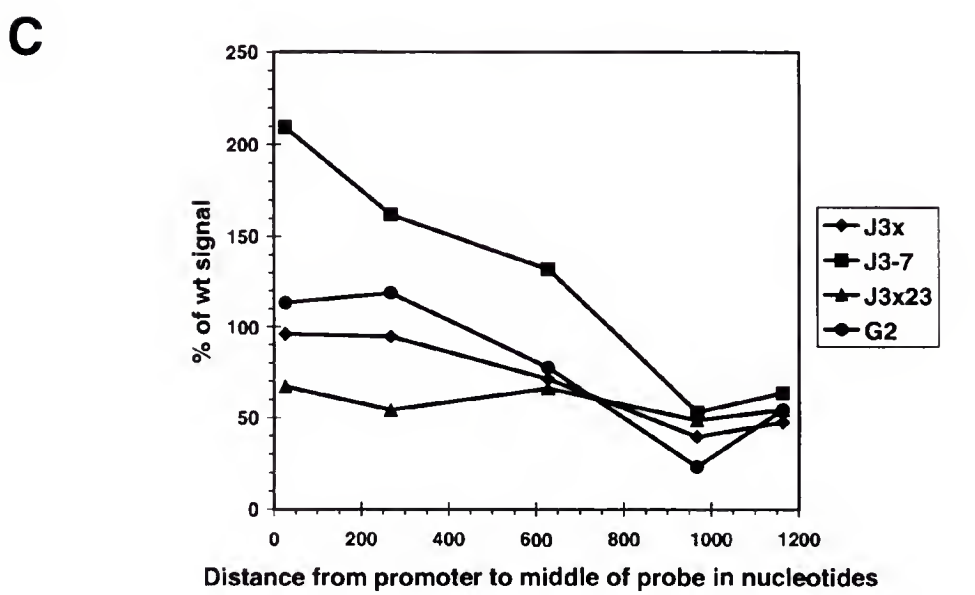

C
250
200
150
100
50
0
% of wt signal
0
200
400
600
800
1000
1200
Distance from promoter to middle of probe in nucleotides
J3x
J3-7
J3x23
G2

Figure 2-7. mRNA 5' cap structure. The 5' ends of eukaryotic and vaccinia transcripts are modified. In vaccinia, a heterodimeric capping enzyme composed of the D1 and D12 proteins performs the first three steps in cap formation by removing the γ-phosphate from the 5' end of a transcript, then by catalyzing the addition of a G residue in an unusual 5'-5' phosphotriester linkage which is resistant to normal ribonucleolytic activity. Last, the capping enzyme methylates the G residue at the N7 position. The J3 protein performs the last step in cap formation which is the methylation of the first transcribed nucleotide at the ribose 2' position, thus converting a cap-0 structure to a cap-1 structure. The arrow indicates the methyl group that is added by J3. The X can represent any base.

The steps in cap formation are as follows:

i) $pppN(pN)_n \rightarrow ppN(pN)_n + P_i$

ii) $pppG + ppN(pN)_n \rightarrow G(5')pppN(pN)_n + PP_i$

iii) $AdoMet + G(5')pppN(pN)_n \rightarrow m^7G(5')pppN(pN)_n + AdoHcy$

iv) $AdoMet + m^7G(5')pppN(pN)_n \rightarrow m^7G(5')pppNm(pN)_n + AdoHcy$

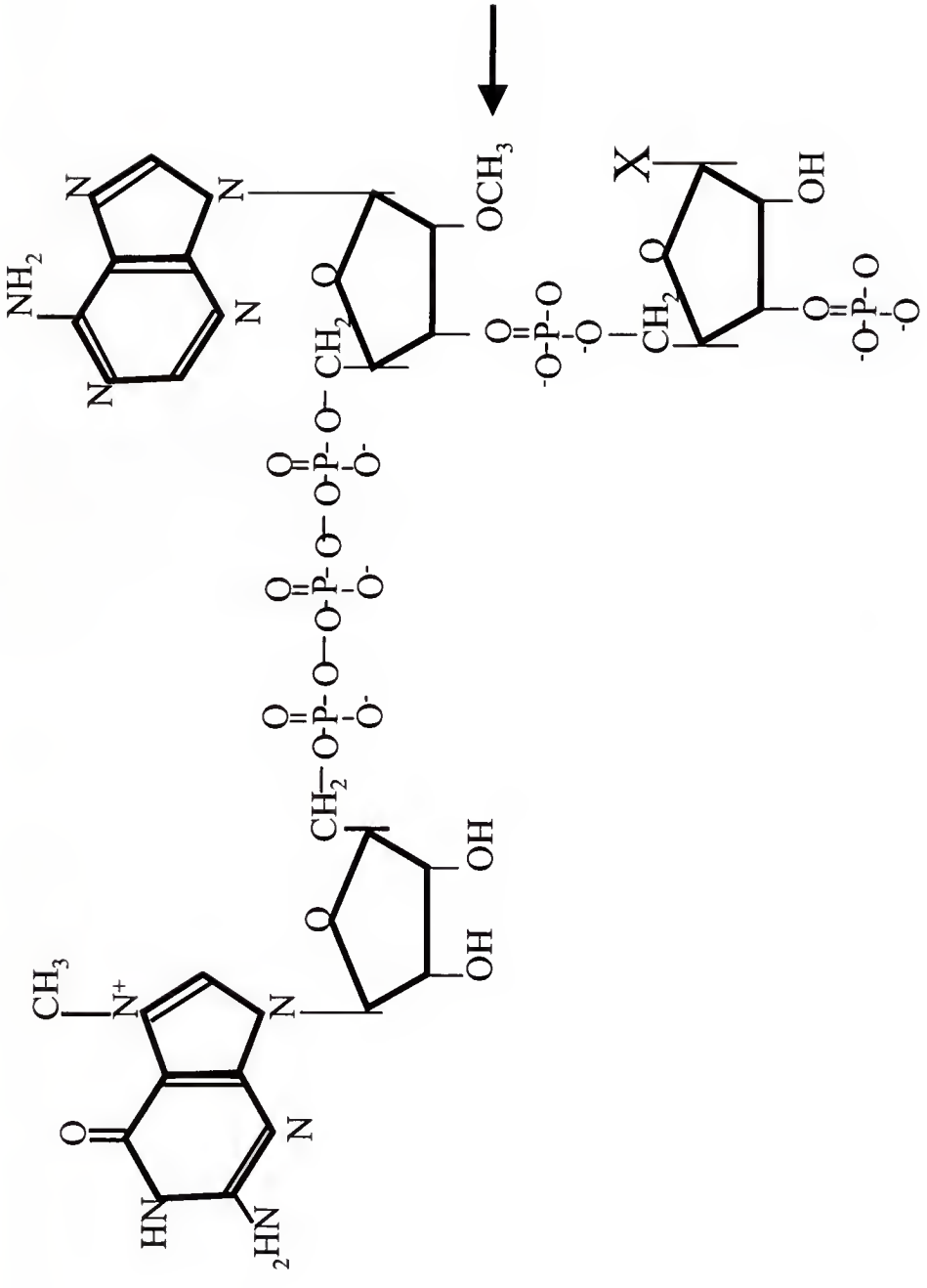
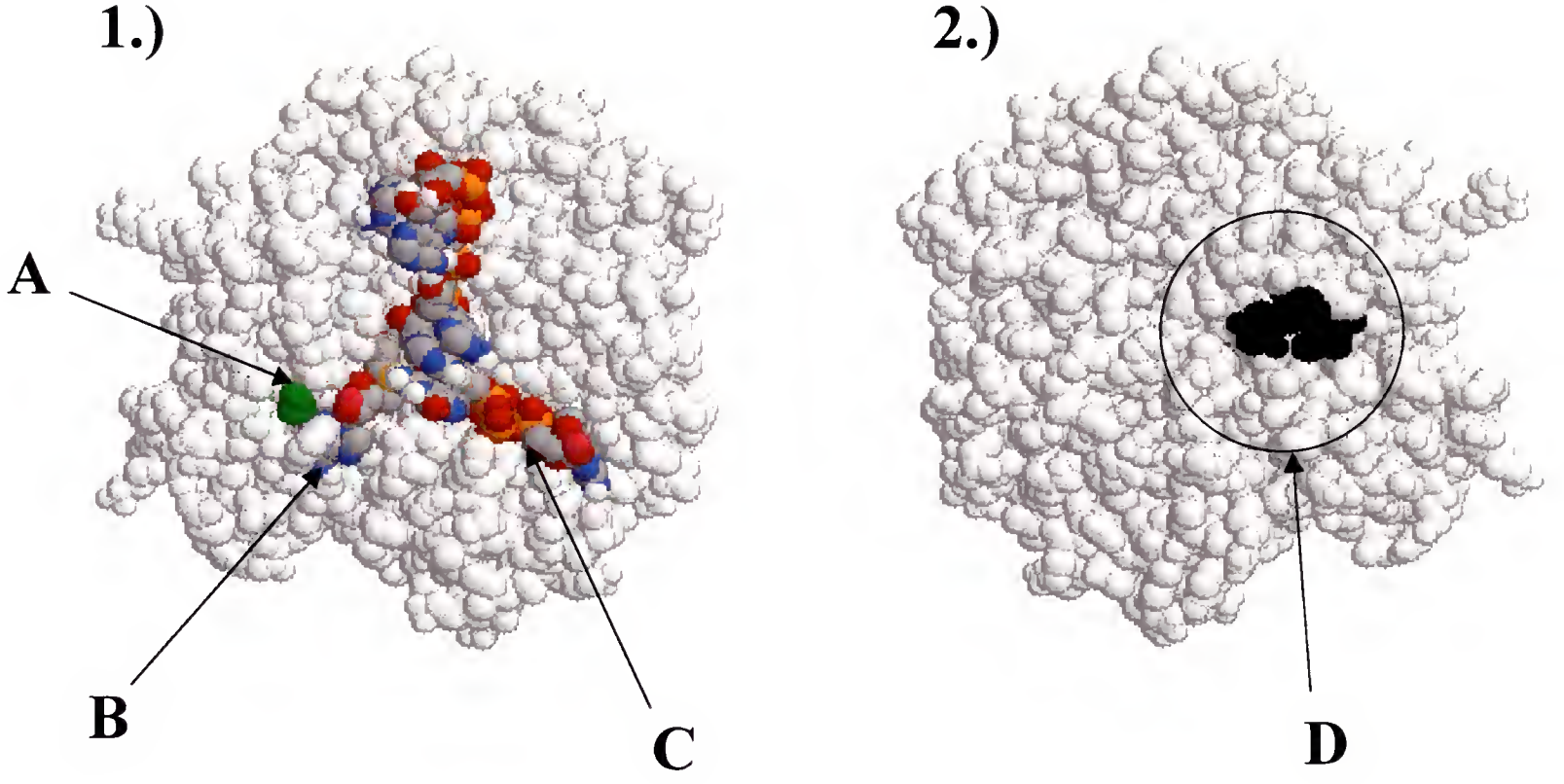


Figure 2-8. The J3 crystal structure. The published J3 crystal structure [PDB (Protein Data Bank) identification number 1AV6] is available through the National Center for Biotechnology Information at http://www.ncbi.nlm.nih.gov and is shown as a spacefilling model. Panel 1 shows the G96D mutation highlighted in green, arrow (A); S-adenosylhomocysetine, the product of the reaction between S-adenosylmethionine and the 2' position of the ribose on the penultimate nucleotide, arrow (B); and a capped 5' end of a mRNA hexamer with the structure m7GpppN$(pN)_5$, arrow (C). Panel 2 shows the opposite face of the molecule rotated 180° backwards relative to panel 1. Arrow (D) indicates residues (highlighted in black) that are important for binding to the E1 poly(A) polymerase catalytic subunit.


1.)
A
B
C
2.)
D

CHAPTER 3
J3 TRANSCRIPTION FACTOR ACTIVITY IS INDEPENDENT FROM ITS OTHER TWO ROLES IN mRNA MODIFICATION

Introduction

As described in chapter 2, the vaccinia J3 protein has been characterized as a postreplicative positive transcription elongation factor. This description is based on two lines of evidence. First, by analogy with the previously described G2 positive transcription elongation factor, J3 mutants have been isolated by two independent genetic selections that implicate a role for J3 in transcription (Latner et al. 2000). Specifically, a J3 point mutant was isolated as an extragenic suppressor of a temperature-sensitive mutation in the A18 transcript release factor. In addition, several J3 null mutants were isolated based upon their dependence for the transcription elongation enhancing drug IBT. Second, it was shown that all of these J3 mutants produce short 3' truncated postreplicative gene transcripts and are thus phenotypically identical to mutants in the G2 gene (Xiang et al. 2000a).

Interestingly, J3 has two previously described roles in mRNA modification. First, it is a (nucleoside-2'-o-)methyltransferase that places a methyl group from S-adenosylmethionine on the 2' position of the first transcribed nucleotide at the 5' end of an mRNA (see figure 2-7 in chapter 2) (Barbosa and Moss 1978a; Barbosa and Moss 1978b). Methylation at the 2' position on the penultimate nucleotide of transcripts is not a novel phenomenon and is a common modification of most eukaryotic mRNA, although

its purpose remains a mystery (Wei and Moss 1975). Second, J3 is the small stimulatory subunit of the heterodimeric viral poly(A) polymerase (Schnierle et al. 1992). The large catalytic subunit of the poly(A) polymerase, E1, will by itself rapidly catalyze the addition of approximately 35 A residues on the 3' end of mRNA at a rate of about 120 residues per minute *in vitro*. After poly(A) tail grows to about 35 bases in length, E1 abruptly ceases rapid elongation of the tail, then very slowly continues to add A residues at a rate of 0.25-1 per minute (Gershon and Moss 1992). Upon addition of J3 to the *in vitro* reaction, E1 returns to a rapid processive mode of elongation until the total length of the poly(A) tail reaches about 200 nucleotides (Gershon 1998). Thus, J3 is involved in modifying both the 5' and 3' ends of mRNAs.

Published site-directed mutagenesis experiments have shown that the (nucleoside-2'-o-)methyltransferase and E1 stimulatory activities are two independent functions of J3 that are genetically separable (Gershon et al 1998). More specifically, these experiments have shown that certain mutations in the J3 protein can precisely abrogate either the (nucleoside-2'-o-)methyltransferase activity or the E1 stimulatory activity as measured *in vitro* with mutant J3 proteins that have been overexpressed and purified from bacteria (Schnierle et al. 1994; Shi et al. 1997). These types of experiments in combination with the more recent observation that J3 is a positive transcription elongation factor beg the following obvious question: Is the J3 positive transcription elongation factor activity somehow functionally linked to one or both of its two other previously described activities, or is the role of J3 as a transcription factor a third independent function of the protein? A direct exploration of this question requires the construction of viruses which contain the J3 mutations that have been shown to disrupt either methyltransferase or

poly(A) polymerase stimulation activity *in vitro*. These mutant viruses must be constructed because an *in vitro* assay to measure the effects of recombinant purified J3 mutant protein on transcription elongation does not exist. Thus, the transcription elongation phenotypes of the mutants must be characterized from virus infected cells. The plasmid clones containing the coding sequence for several previously described J3 site-directed mutants (Schnierle et al. 1994; Shi et al. 1997) have been kindly made available by the laboratory of Dr. Paul Gershon. These clones have been recombined into the virus by transient dominant selection and the mutant viruses have been phenotypically characterized. The results reported here show that the J3 positive transcription elongation factor activity is a third, independent, genetically separable activity of the protein that is otherwise not related to the (nucleoside-2'-o-)methyltransferase or E1 stimulatory activities.

Materials and Methods

Cell Culture, Plaque Assay, One-Step Growth, Protein Pulse-Labeling, RNA Isolation and Northern Analysis

The methods for cell culture, plaque assay, one-step growth, protein pulse-labeling, and northern analysis have all been described in chapter 2.

Transient Dominant Selection

The method for performing transient dominant selection and the plasmids containing the mutated copies of the J3 gene have been previously described (Falkner and Moss 1990; Hassett and Condit 1994; Guan and Dixon 1991; Schnierle et al. 1994; Shi et al. 1997). Briefly, the EcoRI fragment from pBSgpt4 containing the vaccinia 7.5K promoter driving expression of the E.coli gpt cassette was cloned into pGEX/KG

plasmids containing mutated copies of the J3 gene (generously supplied by Dr. Paul Gershon). The sequences of the mutant J3 genes on the recombinant plasmids were verified by sequence analysis prior to transfection. Specifically, the mutant plasmids that contained both the mutant J3 gene and the EcoRI gpt cassette from pBSgpt4 were called pGEX/KG/AS-4/gpt, pGEX/KG/AS-5/gpt, pGEX/KG/AS-6/gpt, pGEX/KG/AS-8/gpt, pGEX/KG/AS-9/gpt, pGEX/KG/AS-15/gpt, pGEX/KG/K175R/gpt, and pGEX/KG/CF3^{c-} /gpt. As a side note, pGEX/KG/CF3^{c-} was constructed by Jody Thompson. A 60mm dish of CV-1 cells was infected at m.o.i. 0.5 with wildtype vaccinia virus (strain WR) diluted in 0.5ml of PBS containing 10mM $MgCl_2$ and 0.01%BSA. Infected cells were incubated at 37°C for 1 hour, then the inoculum was replaced with serum-free 1x F11 media containing 2.5µg/ml mycophenolic acid (from 10mg/ml stock in 100% ethanol), 250µg/ml xanthine (from 10mg/ml stock in 0.1N NaOH), and 15µg/ml hypoxanthine (from 10mg/ml stock in 0.1N NaOH). Cells were returned to 37°C for 2.5 hours. Transfection mixes were in a total volume of 100µl and consisted of 2µg plasmid DNA (prepared by Qiagen mega prep method) and 30µl lipofectin (BRL). Transfection cocktails were added to the infected cells dropwise at 2.5 hours post-infection. Infected-transfected cells were incubated overnight at 37°C and the media was changed to fresh 1x F11 containing 5% fetal calf serum, 2.5µg/ml mycophenolic acid (MPA), 250µg/ml xanthine, and 15µg/ml hypoxanthine. Cells were returned to 37°C until the cytopathic effect was complete (approximately 5 days later). Infected cell lysates were harvested and titered on CV-1 cells pretreated for 12 hours with MPA, xanthine, and hypoxanthine and on CV-1 cells growing in the absence of drugs. Plaques were stained with nutrient agar containing 0.005% neutral red 4 days post-infection as previously described (Condit

and Motyczka 1981; Condit et al. 1983). Plaques growing in the presence of MPA were picked and titered in the presence and absence of 45μM IBT at 31, 37, and 40°C. The MPA resistant plaques that were isolated from cells transfected with the plasmids containing the AS-4, AS-6, AS-9, K175R, and $CF3^{c-}$ alleles were all sensitive to IBT and grew at all three temperatures. The MPA resistant plaques that were isolated from cells transfected with the plasmids containing the AS-5, AS-8, and AS-15 alleles were able to grow both in the presence and absence of IBT at all three temperatures. AS-4, AS-6, AS-9, K175R, and $CF3^{c-}$ candidate mutant plaques that grew at 37°C in the absence of IBT were picked and grown under the same conditions from which they originated. AS-5, AS-8, and AS-15 candidate mutant plaques that grew at 37°C in the presence of IBT were picked and grown under the same conditions from which they originated. The mutant candidates were then screened for the J3 mutations by PCR and RFLP analysis. Candidate J3 mutant viruses were then plaque purified and the entire J3 gene coding region was sequenced from each virus to confirm the presence of the mutations.

In Vivo Poly(A) Tail Measurements

Poly(A) tails were measured essentially as previously described (Birse et al. 1998b; Xiang et al. 2000a). Briefly, 1.5μg of total RNA harvested from cells at 12 hours post-infection was labeled in a 30μl reaction containing 50mM HEPES pH 7.5, 5μM ATP, 10mM $MgCl_2$, 3.3mM DTT, 10% DMSO, 300μg/ml BSA, 20μCi 3'-5'[5'-^{32}P]pCp (3000 Ci/mmol), and 20 units of T_4 RNA ligase (NEB). Reactions were incubated on ice for 18 hours then were digested at 37°C for 4.5 hours by the addition of 50μl of a digest mix containing 40μg tRNA, 100U RNase T_1, 10μg RNase A, 10mM Tris-HCl pH 7.5, and 300mM NaCl. 20μl of a mixture containing 2mg/ml proteinase K, 130mM EDTA,

and 2.5% SDS was added and the reactions were incubated at 37°C for an additional 30 minutes. Reactions were then extracted with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1). The aqueous layer was precipitated by the addition of 60µg glycogen and 260µl 100% ethanol followed by centrifugation for 30 min at 16,100 rcf. Pellets were resuspended in 10µl of a solution containing 95% formamide, 20mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanole. 5µl of each sample was electrophoresed on 8% polyacrylamide 8M urea sequencing gels in 1xTBE at 1700v, 72mA, and 35w with constant power. Gels were fixed in 10% methanol, 10% acetic acid, then were dried and subjected to autoradiography.

Cap Marker Synthesis

Poly(A) RNA was synthesized essentially as previously described (Shuman and Moss 1990) in a 100µl reaction containing 8µg M13 single stranded DNA, 50mM Tris-HCl pH 7.8, 2.5mM DTT, 3mM $MnCl_2$, 1.5mM ATP, 20µM cordycepin triphosphate, 10uCi [α^{32}P]-ATP, 10u E.coli RNA polymerase (USB), and 40u RNasein (Promega). The reaction was incubated at 37°C for 20 hours, then poly(A) RNA was purified by batchwise oligo-dT selection. Briefly, the reaction was made 50mM with respect to EDTA, then was incubated at 70°C for 10 minutes. The reaction was combined with an equal volume of 2x binding buffer [1M NaCl, 20mM Tris-HCl pH 7.8, 2mM EDTA, 1% SDS], and was incubated with 10mg of oligo-dT cellulose (NEB) for 30 minutes. The cellulose-bound poly(A) RNA was pelleted by centrifugation (1 min at 6,000 rcf), then was washed 3 times with 1x binding buffer. Poly(A) RNA was eluted twice with 600µl of elution buffer [10mM Tris-HCl pH 7.8, 1mM EDTA, 0.2% SDS] and was precipitated with 300mM sodium acetate, 40µg glycogen, and an equal volume of isopropanol. The

poly(A) RNA 5' ends were converted to the cap 0 structure essentially as previously described (Shuman and Moss 1990) in a 15µl reaction containing [50mM Tris-HCl pH 7.8, 6mM KCl, 1.2mM $MgCl_2$, 20µCi [α^{32}P]-GTP, 10pmol poly(A) RNA substrate, 150nM vaccinia virus capping enzyme, 1mM S-adenosylmethionine, 40u RNasein (Promega)]. For conversion to cap 1, 280ng of wild type vaccinia virus J3 protein, purified as previously described (Xiang et al. 2000a), was included in the reaction. Capping reactions were incubated for 19 hours at 37°C. Capped poly(A) was once again selected over oligo-dT cellulose and precipitated as described above. The cap-labeled RNA was digested in a reaction containing 10mM Tris-HCl pH 7.8, 50mM sodium acetate, 2mM EDTA, 1µg RNase A (Sigma), 100u RNase T_1 (Boehringer Mannheim), 25u RNase T_2 (BRL), and 1u calf alkaline phosphatase (BRL).

In Vivo Cap Labeling

One 100mm dish of BSC40 cells was infected at m.o.i. = 10 with each virus (except J3-7 which was at m.o.i = 4.8 due to titer constraints). Infected cells were incubated with 1x DME (BRL) and 10% fetal calf serum at 37°C for 10 hours. The media was removed and the infected cells were washed with 4ml of phosphate-free, serum-free 1x DME. Cells were then incubated in the presence of 3.33mCi of [^{32}P]-orthophosphoric acid (285.5Ci/mg) diluted in 4 ml of phosphate-free 1x DME for 2 hours at 37°C. Cells were washed with 4ml of PBS, then total RNA was extracted with an RNeasy kit (Qiagen Inc.) according to the manufacturer's instructions with the following changes: Approximately 10^7 cells were lysed with 1.2ml of RLT buffer + 1% β-mercaptoethanol, and the lysate was spun over 2 separate RNeasy columns. The total RNA from each 100mm dish of infected cells was eluted in a total volume of 200µl water

and made 50mM with respect to EDTA. Samples were heated to 70°C for 5 minutes then poly(A)$^+$ RNA was selected over 10mg oligo-dT resin as described above.

Cap Analysis

Analysis of labeled caps was adapted from a procedure described by Kuge et al (Kuge et al. 1998). Isopropanol precipitated RNA, labeled either in vitro or in vivo, was resuspended in 10µl of digestion mix containing 10mM Tris-HCl pH 7.8, 50mM sodium acetate, 2mM EDTA, 1µg RNase A (Sigma), 100u RNase T_1 (Boehringer Mannheim), 25u RNase T_2 (BRL), and 1u calf alkaline phosphatase (BRL), and incubated for 19 hours at 37°C. To dilute the cap markers while normalizing the salt concentrations relative to in vivo samples, approximately 5000 Cherenkov cpm of RNase treated, in vitro labeled RNA was added to another 10µl of digestion mix before further processing. Digested samples were then concentrated to approximately 1µl in a speed vac, then resuspended in 10µl of [95% formamide, 20mM EDTA, 0.1% bromophenol blue, 0.1% xylene cyanole]. Samples were electrophoresed on a 20% polyacrylamide, 8M urea sequencing gel in 1x TBE at 1700v, 72mA, 33w with constant power. The gel was fixed in a solution of 20% methanol, 20% acetic acid, dried, then subjected to phosphorimager analysis.

Results

Construction of Site-Directed J3 Mutants

To test the hypothesis that the J3 positive transcription elongation factor activity is a third independent activity of the protein, eight J3 site-directed mutant viruses were constructed by transient dominant selection as indicated under materials and methods (Falkner and Moss 1990; Hassett and Condit 1994) using plasmids that were generously

supplied by the Dr. Paul Gershon laboratory. Seven of the eight mutant alleles that were inserted into the virus have been previously described by *in vitro* experiments (Schnierle et al. 1994; Shi et al. 1997). The remaining previously undescribed allele is called K175R. The eight J3 mutant viruses fall into two groups. The first group consists of two viruses, K175R and $CF3^{c-}$, that contain mutations which cause the purified recombinant J3 mutant proteins to be specifically defective *in vitro* for either methyltransferase or poly(A) polymerase stimulation activities respectively. It was hypothesized that if J3 elongation factor activity is a third independent function of the protein, then viruses containing the K175R or $CF3^{c-}$ mutations would produce normal length transcripts, assuming that they maintain their methyltransferase and E1 stimulation defects *in vivo*. Thus, the K175R and $CF3^{c-}$ viruses are two critical controls. The K175R virus contains a missense mutation that converts a lysine to an arginine at codon 175 (unpublished). Based on the crystal structure, it is known that the wildtype lysine residue at position 175 lies in the active site of the methyltransferase and is important for binding of the methyl donor S-adenosylmethionine (SAM). Exchange of the lysine for an arginine thus compromises the ability of the mutant to bind SAM and the purified recombinant protein has been shown to be defective for methyltransferase activity *in vitro* (Paul Gershon personal communication). The $CF3^{c-}$ virus contains two missense mutations that convert the histidine at codon 56 to an arginine (H56R) and convert the isoleucine at position 58 to a serine (I58S) (Shi et al. 1997). These mutations are located on a surface region of the protein that is important for binding the E1 poly(A) polymerase and the purified recombinant protein is thus defective for poly(A) tail synthesis *in vitro* (Shi et al. 1997). Of note, the $CF3^{-}$ virus described here lacks two mutations that were originally described

in the initial characterization of the purified recombinant protein. Specifically, the previously described recombinant $CF3^{c-}$ protein contains two additional missense mutations that convert cysteine 178 and cysteine 272 to serine residues and were shown not to affect the assayable properties of the protein (Shi et al. 1997; and P. Gershon, personal communication). The remaining viruses fall into a second group which contain charge-to-alanine scanning (AS) mutations that are located at or near the surface of the protein, based on crystal structure observations (Hodel et al. 1996a). It has been previously shown that the purified recombinant J3 proteins containing each of these mutations are normal for methyltransferase and E1 stimulation activities *in vitro* (Schnierle et al. 1994). It was hypothesized that some of these mutations might disrupt a region of J3 that is important for binding to an additional putative transcription factor and thus produce a specific defect in transcription elongation while maintaining normal methyltransferase and E1 stimulation activities. Therefore, construction of these viruses could potentially map on the surface of J3 a region that might be responsible for stimulating transcription. The charge-to-alanine scanning mutant viruses that were constructed are called AS-4, AS-5, AS-6, AS-8, AS-9, and AS-15. The specific mutations contained in each of these viruses are listed in table 3-1. Briefly however, the following residues were mutated to alanine in each virus as indicated (the numbers signify the codon position): (AS-4) arginine 79, aspartate 80; (AS-5) arginine 97, histidine 98, histidine 99; (AS-6) arginine 107, aspartate 108; (AS-8) arginine 122, lysine 125, lysine 126; (AS-9) histidine 129, lysine 132; (AS-15) arginine 220, arginine 223. As described under Materials and Methods, each of the viruses indicated above were constructed by transient dominant selection, isolated by screening plaques with PCR

and restriction fragment length polymorphism analysis, plaque purified, and the mutations in the J3 genes were verified by sequence analysis.

Plaque Phenotypes of Site-Directed Mutants

Previous experience with other G2 and J3 mutants has shown that the response of a virus to IBT is a good indicator of the mRNA synthesis phenotype of the virus (Black and Condit 1996; Latner et al. 2000; Xiang et al. 2000a). To more specifically illustrate, wildtype virus is IBT sensitive (IBT^S) and makes normal length transcripts while IBT dependent (IBT^D) G2 and J3 mutants make short 3' truncated transcripts. Therefore, each of the site-directed mutants were examined in a plaque assay for their response to IBT. Figure 3-1 demonstrates that as controls, the wt virus is IBT sensitive while the previously characterized *J3x* and *J3-7* viruses are IBT dependent. As described in chapter 2, the *J3x* virus is somewhat leaky at 37°C as attested by the presence of very small, barely visible plaques in the absence of drug that are equal in number to the plaques formed in the presence of drug. *J3-7*, as described in chapter 2 and in (Latner et al. 2000), has a tight IBT^D phenotype but forms very small plaques, consistent with the fact that it is difficult to grow to high titer. Importantly, the $CF3^{c-}$ and K175R viruses are both IBT sensitive like wt, suggesting that they may have normal transcription phenotypes. Interestingly, the $CF3^{c-}$ virus consistently produced slightly smaller-than-wt plaques, suggesting that it may be somewhat defective for growth. The AS-4, AS-6, and AS-9 viruses are all IBT sensitive, suggesting that that they also contain no transcription defects. Interestingly, the AS-5, AS-8, and AS-15 viruses are all IBT resistant, preliminarily suggesting that they may have some defect in transcription elongation. With IBT resistance conceptually being an intermediate phenotype on the continuum

somewhere between sensitivity and dependence, it was hypothesized that AS-5, AS-8, and AS-15 may produce transcripts that are of an intermediate length. That is, transcripts produced by an IBT resistant mutant may not be as long as the normal wt length transcripts, but not as short as the transcripts produced by the IBT dependent viruses *J3x* and *J3-7*.

One-Step Growth Analysis

To support the plaque phenotype data, one-step growth analysis was performed on each mutant (figure 3-2). Briefly, 60-mm dishes of confluent BSC40 cells were infected with each virus at m.o.i. = 3.3 and were incubated at 37°C in the presence or absence of IBT. Infected cell lysates were then harvested at 0 hr or 72 hr post-infection and were plaque titrated under permissive conditions for each virus. Specifically, all of the lysates with the exceptions of *J3x* and *J3-7* were titered in the absence of IBT at 37°C. The *J3x* and *J3-7* lysates were titered in the presence of IBT at 37°C. The yield from each infection is summarized in figure 2. By 72 hr postinfection, wt virus produced approximately 77 pfu/cell, significantly higher than the T_0 background, in the absence of IBT and was inhibited for growth in the presence of IBT. *J3x* produced only background levels of progeny in the absence of IBT and approximately 0.8 pfu/cell in the presence of IBT, which is approximately 10-fold higher than background. This confirms that *J3x* is IBT dependent and is difficult to grow to high titer. *J3-7* produced background levels of progeny in the absence of IBT and slightly more in the presence of drug, confirming that although *J3-7* is IBT dependent, it is significantly crippled and is very difficult to grow to a high titer. AS-5 produced between 13 and 17 pfu/cell both in the presence and absence of IBT, demonstrating that it is clearly resistant. AS-8 produced 6 pfu/cell in the absence

of drug and 3 pfu/cell in the presence, showing that while it is also IBT resistant, it may be weakly inhibited by the drug. AS-15 produced between 22 and 32 pfu/cell in the presence and absence of IBT, showing that it is strongly IBT resistant. The $CF3^{c-}$ virus produced approximately 31 pfu/cell in the absence of IBT and background levels of progeny in the absence of drug. Thus, $CF3^{c-}$ is IBT sensitive like wt, but relative to wt may be somewhat hindered for growth. This data supports the observation that $CF3^{c-}$ consistently produces slightly smaller plaques than wt. Previous one-step growth experiments have consistently shown, however, that $CF3^{c-}$ eventually reaches wildtype titer (data not shown). The K175R, AS-4, AS-6, and AS-9 viruses all produced between 106 and 125 pfu/cell in the absence of IBT and only 0.02-0.41 pfu/cell in the presence of IBT. Therefore, they are all IBT sensitive and grow to wt titer. In summary, these results support the plaque phenotypes shown in figure 3-1 and show that the $CF3^{c-}$ and K175R control viruses and three charge-to-alanine scanning mutants are IBT sensitive. In addition, the one-step growth data confirm that the AS-5, AS-8, and AS-15 viruses are all IBT resistant.

Viral Protein Synthesis

As an initial test of the prediction that the J3 site-directed mutant viruses would have a transcription phenotype that is consistent with their response to IBT, the pattern of viral gene expression was examined by metabolically labeling proteins in infected cells with [^{35}S]methionine at various times postinfection. The total infected cell protein was then analyzed by SDS-PAGE and autoradiography. It was determined that the protein synthesis profile of each virus can be placed into one of three different phenotypic categories that directly correlate with the response of the virus to IBT. Shown in figure 3-3 are autoradiograms that are representative examples of each of the phenotypes. The

first phenotypic category is composed of the IBT sensitive viruses and is illustrated by the wt and K175R viruses. In the absence of IBT, the wt virus shuts off host protein synthesis between 3 and 6 hours postinfection, early viral protein synthesis begins between 0 and 3 hrs and subsequently decays, and synthesis of intermediate and late virus proteins begins between 3 and 6 hrs and is maintained throughout infection. This profile contrasts with the wt infection in the presence of IBT, where the synthesis and shut off of early proteins is normal, late protein synthesis begins, and all protein synthesis decays by 9 hrs postinfection. This observation is consistent with previously published results (Cooper et al. 1979; Black and Condit 1996) and is largely attributed to induction of the 2-5A RNA degradation pathway which is turned on in IBT-treated wt infections (Pacha et al. 1990; Meis and Condit 1991; Black and Condit 1996). Although the mechanism of IBT action is unknown, it appears to stimulate transcription elongation (Pacha and Condit 1985; Pacha et al. 1990; Bayliss and Condit 1993; Simpson and Condit 1994; Black and Condit 1996). Vaccinia genes are closely spaced and the transcriptional units are often pointing towards one another (Moss 1990b; Moss 1996). When transcription elongation is enhanced by the addition of IBT to an infection, the excessively long transcripts that are generated from opposing promoters can hybridize to one-another. Thus IBT treatment leads to the formation of double-stranded RNA and induces the double-stranded RNA dependent 2-5A RNA degradation pathway (Woodson and Joklik 1965; Cooper et al. 1979; Cohrs et al. 1989). Protein synthesis then ceases as a pleiotropic effect of RNA degradation. All of the IBT sensitive viruses, which include $CF3^{c-}$, K175R, AS-4, AS-6, and AS-9 show a protein synthesis profile that is similar to the wt virus. An additional representative example is shown by K175R in figure 3-3.

The second phenotypic category is composed of the IBT dependent viruses and is represented by *J3x*. When the *J3x* mutant is grown in the absence of IBT, it displays normal shut-off of host protein synthesis by 6 hrs postinfection and has normal early protein synthesis that decays between 3 and 6 hrs postinfection (Xiang et al. 2000a). Beginning at 6 hrs and maintained throughout infection, *J3x* produces normal amounts small intermediate and late proteins, but produces reduced quantities of large intermediate and late proteins. As described in chapter 2, this specific defect in protein synthesis is due to the fact that the *J3x* postreplicative transcripts are 3' truncated and are too short to encode full length large proteins (Xiang et al. 2000a). In the presence of IBT, *J3x* shows normal shut-off of host protein synthesis and early protein synthesis occurs and decays normally. Interestingly, in the presence of IBT, *J3x* produces roughly equivalent amounts of both large and small proteins at late times during infection, indicating that the defect in *J3x* transcription has somehow been corrected by the addition of drug. Thus the profile of *J3x* protein synthesis in the presence of IBT is comparable to the wt profile in the absence of IBT in terms of the kinetics of synthesis and the size and quantity of proteins produced.

The third phenotypic category is composed of the IBT resistant viruses AS-5, AS-8, and AS-15 and is represented here by AS-5. Both in the presence and absence of IBT, the pattern of AS-5 protein synthesis appears identical to the pattern protein synthesis demonstrated by the wt virus growing in the absence of drug. Thus regardless of IBT treatment, AS-5 has normal shut-off of host synthesis, normal early protein synthesis and shut-off, and normal intermediate and late protein synthesis that persists throughout infection. This result suggests that the J3 mutations in the AS-5, AS-8, and AS-15

viruses may have some moderate effect on transcription elongation. Hypothetically, the IBT resistant viruses may produce transcripts that are of an intermediate length, that is not as long as wt and not as short as those produced by the IBT dependent mutants. Thus, unlike the IBT dependent mutants, resistant viruses would not require IBT to synthesize full length transcripts that could code for the large late proteins, and unlike the sensitive viruses, the presence of IBT would not induce excessive elongation that would lead to the accumulation of double-stranded RNA to levels that would induce the 2-5A pathway.

Viral mRNA Synthesis

A crucial step in determining if the J3 transcription elongation factor activity is linked to the other two functions of the protein is to examine the mRNAs synthesized by each site-directed mutant. Of particular importance are the $CF3^{c-}$ and K175R viruses. If these two viruses synthesize normal length mRNA but retain their defects in methyltransferase and E1 stimulation activity *in vivo*, then the J3 elongation factor activity must be a third independent function of the protein. Similarly, additional evidence unlinking the three activities might be provided if the IBT resistant mutants synthesize mRNAs that are somewhat shorter than wt and if they retain *in vivo* methyltransferase and E1 stimulation functions. Thus mRNAs produced by each virus were examined by northern blot analysis. Briefly, total RNA was harvested from infected cells at various times postinfection, electrophoresed on formaldehyde agarose gels, transferred to nylon membranes, and hybridized with standard early (gene C11), intermediate (gene G8), and late (gene F17) gene riboprobes. Figure 3-4 panel A shows RNA harvested at 3 hrs postinfection and probed with an early gene riboprobe. As previously described in chapter 2 and in (Xiang et al. 2000a), the early wt transcripts are

characteristically homogeneous in length due to the *cis*-acting termination signal. Importantly, all of the site-directed mutant viruses produce early transcripts that are essentially identical in length to those produced by wt, indicting they have normal early mRNA synthesis. When the infections are performed in the presence of IBT (panel B), the early transcripts still appear normal in all cases, indicating that IBT has no effect on early transcription.

When RNA harvested at 9 hrs postinfection from cells infected in the absence of IBT is probed with the intermediate gene riboprobe (figure 3-4 panel C), a characteristic heterogeneous population of transcripts is detected in all cases. Consistent with the results shown in chapter 2, the wt virus produces a normal distribution of intermediate transcripts and the IBT dependent mutants, *J3x* and *J3-7*, produce a population of transcripts that is on average shorter than those produced by the wt virus. Importantly, the $CF3^{c-}$ and K175R viruses both produce intermediate transcripts that are of wt length, indicating that they do not have a defect in intermediate gene transcription. The remaining IBT sensitive viruses, AS-4, AS-6, and AS-9, also produce intermediate transcripts that are of a normal length. Interestingly, the intermediate transcripts produced by the IBT resistant mutants AS-5, AS-8, and AS-15, are similar to the transcripts produced by the other IBT sensitive viruses. This suggests that if the AS-5, AS-8, and AS-15 viruses have a defect in elongation as initially hypothesized, it is a subtle defect and may require a more sensitive assay for better characterization. When RNA harvested at 9 hrs postinfection from cells infected in the presence of IBT is probed with the intermediate gene riboprobe, some interesting results are obtained (figure 3-4 panel D). First, the ribosomal RNA bands that can normally be clearly seen on the

ethidium-bromide stained formaldehyde agarose gels prior to nylon-transfer are degraded specifically from cells that were infected with IBT sensitive viruses. In contrast, the ribosomal RNAs remain intact from cells infected with either IBT dependent or IBT resistant mutants (data not shown). This observed RNA degradation is consistent with the previously described induction of the 2-5A pathway in cells that are infected with IBT sensitive viruses in the presence of IBT. When probed with the intermediate gene riboprobe, the wt (+IBT) sample shows a distribution of transcripts that is different from the distribution of the wt (-IBT) sample. Specifically, the size distribution of transcripts produced in the presence of IBT is between approximately 0.1 to 5-kb while the transcripts detected in the absence of IBT are between approximately 1.4 to 5-kb. The appearance of these small transcripts in the presence of IBT is also consistent with RNA degradation. The same distribution is observed from each IBT sensitive virus sample, including $CF3^{c-}$, K175R, AS-4, AS-6, and AS-9, demonstrating that they, like wt, experience RNA degradation in the presence of IBT. Interestingly, the IBT dependent viruses *J3x* and *J3-7* both appear to produce transcripts in the presence of IBT that resemble the lengths of the transcripts produced from the wt virus in the absence of IBT, that is they are mostly between 1.4 and 7-kb in length. In comparison to the transcripts that *J3x* and *J3-7* produce in the absence of drug, the transcripts they produce in the presence of drug appear longer on average. This suggests that IBT is compensating for the defects that these viruses have in transcription. In the presence of IBT, the IBT resistant viruses AS-5, AS-8, and AS-15 all appear to produce intermediate transcripts that are of similar length to the IBT dependent viruses *J3x* and *J3-7*.

Figure 3-4 panel E shows RNA harvested at 12 hrs postinfection and probed with a late gene riboprobe. Once again, wt virus produces a normal heterogeneous distribution of transcripts with a strong homogeneous band of 1.4-kb in length that is consistent with previous observations (chapter 2). The *J3x* and *J3-7* IBT dependent viruses produce transcripts that are on average, shorter than those produced by the wt virus. The remaining IBT sensitive and IBT resistant viruses produce transcripts in the absence of drug that are not distinctly different from the wt transcripts. Figure 3-4 panel F shows RNA harvested at 12 hrs postinfection from cells infected in the presence of IBT and probed with a late probe. By 12 hrs postinfection, it appears that wt and the other IBT sensitive viruses have reduced production of late transcripts relative to the amount of transcripts produced in the absence of IBT. This observation is probably a result of the degradation of intermediate transcripts, which encode late transcription factors. The *J3x* and *J3-7* IBT^{d} viruses produce late transcripts, that by analogy with the intermediate transcripts produced in the presence of drug, are slightly longer than the transcripts produced in the absence of drug. This once again suggests that IBT is compensating for the transcription defects in these viruses, allowing them to synthesize normal length transcripts. The IBT resistant viruses AS-5, AS-8, and AS-15, unlike the sensitive viruses, produce late transcripts in the presence of drug. In addition, the transcripts they produce in the presence of drug are that are essentially of the same length as the transcripts they produce in the absence of drug.

In Vivo Poly(A) Polymerase Stimulatory Activity

After determining that the CF3^{c-} and K175R viruses are wt for RNA synthesis, the next step in determining if the J3 transcription elongation activity is linked to the other two functions was to examine each of the viruses for *in vivo* poly(A) polymerase

stimulatory activity. Briefly, cells were infected with each virus in the absence of IBT and total RNA was harvested at 12 hrs postinfection. The RNAs were then labeled at the 3' end with [^{32}P]pCp (cytidine 3', 5' –bisphosphate) by T_4 RNA ligase and were digested with RNases A and T_1. The RNase resistant labeled poly(A) tails were then analyzed on a polyacrylamide sequencing gel and the autoradiogram is shown in figure 3-5. As previously described (Xiang et al. 2000a), wt virus synthesizes poly(A) tails that are between 0 and 150 nucleotides in length. The *J3x*, K175R, AS-4, AS-5, AS-6, AS-8, and AS-15 viruses all synthesize poly(A) tails that are essentially identical to those produced by the wt virus, indicating that they have normal poly(A) polymerase stimulatory activity. In contrast, the *J3-7* IBT dependent virus, which is a J3 null mutant, is defective for poly(A) tail synthesis as previously described. Most of the poly(A) tails synthesized by *J3-7* are less than 50 nucleotides in length, which is consistent with this virus lacking the poly(A) polymerase stimulatory function. Notably, the CF3^{c-} virus, which contains a mutation that has been previously described as disrupting poly(A) polymerase stimulatory activity *in vitro* (Shi et al. 1997), appears to have a defect in poly(A) tail synthesis *in vivo*, like *J3-7*. Specifically, the poly(A) tails on the CF3^{c-} transcripts are also predominantly less than 50 nucleotides in length. Interestingly, it seems as if overall there are fewer poly(A) tails in the CF3^{c-} sample. This observation is supported by the phosphorimage analysis that is graphically displayed in figure 3-6. Briefly, the graph shows a comparison of the relative amounts of radioactivity detected in each sample (y-axis) as a function of the distance from the top of the gel (x-axis). This apparent reduction in the abundance of poly(A) tails could be due to three different things: 1) a reduction in overall abundance of transcripts produced by CF3^{c-}; 2) a failure to

efficiently recover small poly(A) tails in the labeling procedure; or 3) $CF3^{c-}$ is simply very inefficient at polyadenylating transcripts. The possibility that there are fewer $CF3^{c-}$ transcripts is unlikely based on the northern blots in figure 3-4 which show a roughly equivalent number of transcripts relative to the other viruses including wt. Several modifications to the poly(A) tail measurement protocol were made in attempt to enhance recovery of the small poly(A) tails in all samples but were unsuccessful. Thus, it seems most likely that $CF3^{c-}$ is simply very inefficient at polyadenylation. Thus the *J3-7* and $CF3^{c-}$ viruses appear similar to each other and different than wt with respect to poly(A) tail synthesis. Interestingly, the AS-9 virus consistently appeared to have a very slight defect in E1 stimulatory function also, with most of the tails being less than 100 bases in length as detected in several independent experiments. However, the AS-9 defect is difficult to visualize from the experiment shown in figure 3-5.

In Vivo 2'-o-Methyltransferase Activity

The final step in determining if the J3 transcription elongation activity is linked to the other two J3 functions was to examine the in vivo 2'-o-methyltransferase activity of each virus. Briefly, 100-mm dishes of BSC40 cells were infected with each virus and were incubated at 37°C for 10 hours. The infected cells were then labeled with 3.33 mCi of [^{32}P]orthophosphoric acid for 2 hours. Poly(A)$^+$-RNA was then purified and digested with RNases A, T_1, T_2, and calf intestinal alkaline phosphatase (CIAP). The cap structures on the 5' ends of mRNAs are resistant to ribonucleolytic digestion due to the 5'-5' nature of the linkage between the 7-methyl G residue and the first transcribed nucleotide. In addition, if the first transcribed nucleotide is methylated at the 2' position, the first phosphodiester bond at the 5' end is also protected from ribonucleolytic

cleavage. After digestion, an mRNA is thus reduced to nucleosides, phosphates, and either a di-nucleotide or tri-nucleotide that represents the cap structure. The digestion products were electrophoresed on a 20% polyacrylamide 8M urea sequencing gel which was dried and subjected to phosphorimager analysis (figure 3-7). The results show that wt viral transcripts have a cap-1 structure as predicted. Interestingly, the J3x mutant transcripts also have a cap-1 structure. This result is quite surprising because the purified recombinant *J3x* protein has been shown to be defective for methyltransferase activity in an in vitro assay (Xiang et al. 2000a). Although it is possible that a cellular methyltransferase could mask a defect in the J3 methyltransferase activity *in vivo*, it is very unlikely given the results of two controls. Specifically, examination of the caps produced by the J3-7 null mutant and the K175R point mutant shows that transcripts produced by both viruses have cap-0 structures in vivo. The J3-7 and K175R methyltransferase results are consistent with previous descriptions of both mutants. More specifically, J3-7 is a null mutant and produces no detectable J3 protein during an infection (Latner et al. 2000). Thus, it would not be expected to provide 2'-o-methyltransferase activity to an infection. The K175R mutant has been shown to be defective for 2'-o-methyltransferase activity in vitro due its missense mutation which is thought to inhibit SAM binding to the active site of the methyltransferase domain (unpublished results; Dr. Paul Gershon, personal communication). The remaining viruses, $CF3^{c-}$, AS-4, AS-5, AS-6, AS-8, AS-9, and AS-15 all produce cap-1 in vivo.

Discussion

Knowing that the vaccinia virus J3 protein has two previously characterized biochemically independent roles in modifying both the 5' and 3' ends of mRNA

(Gershon et al. 1998), the recent realization that J3 also has a positive transcription elongation factor activity raised the following question: Is the J3 transcription factor activity related to the two other mRNA modification activities or is it a third independent function of the protein? To directly address this question, several site-directed J3 mutant viruses were constructed in attempt to create viruses that are specifically defective for each of the three J3 functions. Rather than simply assaying the previously described recombinant purified mutant proteins, actual viruses containing the site-directed mutations had to be constructed because an *in vitro* assay for the J3 transcription elongation factor activity does not exist. Eight site-directed J3 mutant viruses were constructed by transient dominant selection. Two of the viruses, called K175R and $CF3^{c-}$, contained mutations that were known to disrupt either the methyltransferase or poly(A) polymerase stimulatory functions respectively (as assayed *in vitro*) (Shi et al. 1997; and Paul Gershon, personal communication). The six remaining viruses contained charge-to-alanine scanning (AS) mutations that are located at or near the surface of the J3 protein based on crystal structure data (Hodel et al. 1996a). These mutants were constructed in attempt to create viruses that are specifically defective for transcription elongation and thus map onto the surface of J3 a region that is important for binding additional unidentified transcription factors. Subsequently, each of the mutant viruses were biochemically assayed *in vivo* for the three J3 activities. The location of each J3 mutation and the results of the biochemical characterization are summarized in Table 3-1.

The results provide several lines of evidence showing that the J3 transcription elongation factor activity is a third independent function of the protein that is not related to the 2'-o-methyltransferase or E1 poly(A) polymerase stimulatory activities. First,

based on northern analysis and supported by the protein synthesis profiles and plaque phenotypes, it appears that the $CF3^{c-}$ and K175R viruses have normal transcription throughout infection that is indistinguishable from the wt virus. It was also shown that the $CF3^{c-}$ virus is specifically defective for *in vivo* poly(A) tail synthesis and the K175R virus is specifically defective for *in vivo* 2'-o-methyltransferase activity. Thus the sole defect in $CF3^{c-}$ is for E1 poly(A) polymerase stimulation and the sole defect in K175R is for 2'-o-methyltransferase activity. Second, it was surprisingly revealed that the only defect in the *J3x* point mutant is for transcription elongation activity. Although previous evidence has demonstrated that the recombinant purified *J3x* protein does not have 2'-o-methytransferase activity *in vitro*, the methyltransferase activity is clearly present *in vivo*. Therefore, three different viruses, $CF3^{c-}$, K175R, and *J3x*, each contain a different defect in one of the three J3 functions and together demonstrate that the positive transcription elongation activity of J3 is a third independent function of the protein that can be genetically separated from the other two activities. Moreover, these results show that the J3 transcription elongation activity is the only essential J3 activity for virus growth in tissue culture.

The fact that the $CF3^{c-}$ virus has normal transcription and is defective for E1 binding *in vitro* suggests that the J3 transcription activity does not require that J3 and E1 form a heterodimer. In addition, the results suggest that a particular surface area of J3 may be important for mediating its positive transcription elongation factor activity. Evidence for this comes from the characterization of the *J3x* point mutant and the charge-to-alanine scanning mutants. Shown in figure 3-8 is the J3 crystal structure with particular residues highlighted according to their effects on transcription. The projection

labeled "front" shows the methyltransferase active site bound to S-adenosylhomocysteine (to the left of the arrow), which is the product of the reaction between the methyl donor S-adenosylmethionine and the 2' hydroxyl of the penultimate nucleotide. In addition, the cap 7-methyl G triphosphate is also displayed in the active site to the right of the arrow. The residues that are important for binding the E1 poly(A) polymerase are circled. The residue highlighted in purple is the glycine residue at codon 96 which is mutated to aspartate in the IBT dependent point mutant *J3x*. Realizing that G96 is located near the methyltransferase active site helps to explain why the *J3x* purified protein is defective for *in vitro* methyltransferase activity, but is active *in vivo*. Glycine 96 is very near the active site of the methyltransferase domain and provides a kink in the alpha carbon backbone between two charged residues, aspartate 95 and arginine 97, which have been shown to form hydrogen bonds with the methyl donor SAM. Replacement of glycine 96 with an additional aspartate would therefore be likely to disrupt SAM binding under *in vitro* conditions. However, the *in vivo* conditions could possibly be more conducive for SAM binding which may be facilitated by pH, salt concentrations, or the binding of J3 to other protein factors like the E1 poly(A) polymerase, the NPH-I enzyme, the A18 transcript release factor, or the H4 subunit of the RNA polymerase, all of which have been shown to bind J3 and are discussed in more detail in chapter 4 (Mohamed et al. 2001; and Ed Niles, personal communications). Regardless, the only *in vivo* defect in *J3x* is for transcription elongation activity.

The charge-to-alanine scanning mutations in the IBT resistant viruses AS-5, AS-8, and AS-15 are highlighted in green on the crystal structure shown in figure 3-8. It was initially hypothesized that in the absence of IBT, these mutants would produce transcripts

that are 3' truncated like those produced by the IBT dependent viruses, but that they would not be as dramatically truncated. However, the transcripts produced by the IBT resistant viruses in the absence of drug do not appear significantly different in length from those produced by the wt virus (figure 3-4). One possibility could be that the effects of these mutations on elongation in the absence of drug are so subtle that they cannot be precisely detected by conventional methods. A more dramatic result is obtained when the transcripts produced by these viruses in the presence of IBT are examined. As detected by both the intermediate and late gene probe in the northern analysis, the resistant viruses produce relatively normal length transcripts in the presence of drug. In comparison to the degraded wt transcripts, this clearly demonstrates that they have an altered transcription phenotype relative to the wt virus. Like the *J3x* mutation, these charge-to-alanine scanning mutations could hypothetically alter J3 binding to an additional transcription factor and thus give rise to the aberrant transcription phenotype. This possibility will be discussed in more detail in chapter 4. The residues highlighted in black in figure 3-8 show the mutations in the IBT sensitive charge-to-alanine scanning mutants AS-4, AS-6, and AS-8 and the mutations in the $CF3^{c-}$ and K175R viruses. The K175R mutation is buried under the methyltransferase active site and is indicated by the arrow in the "front" projection. These mutations had no observable effect on transcription.

To briefly summarize the data as it relates to the crystal structure, the locations of the IBT dependent (purple) and resistant mutations (green) shown in figure 3-8 reveal that the surface exposed residues along the left side of the molecule may be important for stimulating elongation. Although it is possible that the mutations alter the protein

conformation at some distance from the precise location of the amino acid substitution, this is unlikely because all of these J3 mutants retain both 2'-o-methyltransferase and E1 poly(A) polymerase stimulatory activities, suggesting that protein conformation is globally conserved. A mechanistic explanation of how these residues could be involved in transcription is proposed in chapter 4.

Table 3-1: Summary of J3 Site-Directed Mutants

		in vitro		in vivo			
Virus	Mutation	MTase	E1-S	MTase	E1-S	Elongation	IBT
wt	NA	+	+	+	+	+	S
J3x	G96D	-	+	+	+	-	D
J3-7	[a]*fs* 49; 58aa	NA	NA	-	-	-	D
CF3[c-]	H56R, I58S	+	-	+	-	+	S
K175R	K175R	-	+	-	+	+	S
AS-4	R79, D80 –A	+	+	+	+	+	S
AS-5	R97, H98, H99 -A	+	+	+	+	+/-[b]	R
AS-6	R107, D108 –A	+	+	+	+	+	S
AS-8	R122, K125, K126 -A	+	+	+	+	+/-[b]	R
AS-9	H129, K132 –A	+	+	+	+	+	S
AS-15	R220, R223 -A	+	+	+	+	+/-[b]	R

[a] *fs* indicates the *J3-7* mutant contains a single nucleotide deletion in codon 49 that produces a frameshift truncation of the protein 58 amino acids from the N-terminus. Each of the remaining missense mutations are indicated as the wt amino acid, location of the codon which is mutated, and the mutant residue. "+" indicates normal wt function and "-" indicates defective or mutant function. S indicates IBT sensitivity; R indicates IBT resistance; D indicates IBT dependence. MTase indicates (nucleoside-2'-o) methyltransferase activity; E1-S indicates E1 poly(A) polymerase stimulatory activity. Elongation indicates length of intermediate and late gene transcripts relative to wt. +/-? indicates transcription may be defective.

[b] +/- indicates that the IBT resistant mutants AS-5, AS-8, and AS-15 produce normal length transcripts in the absence of IBT like wildtype, but they have a transcription phenotype in the presence of IBT that is different from the wildtype virus. Please refer to the text for details.

Figure 3-1. Plaque phenotypes of mutant viruses. Confluent monolayers of BSC40 cells in 60-mm dishes were infected with an appropriate dilution of virus and incubated at 37°C in the presence or absence of IBT under an agar overlay for 4 days. Dishes were stained overnight with a second neutral red containing agar overlay. Agar was then removed and cells were stained with crystal violet. The mutant used for infection is indicated at the top of each column. The presence or absence of IBT during the infection is indicated at the left of each row.

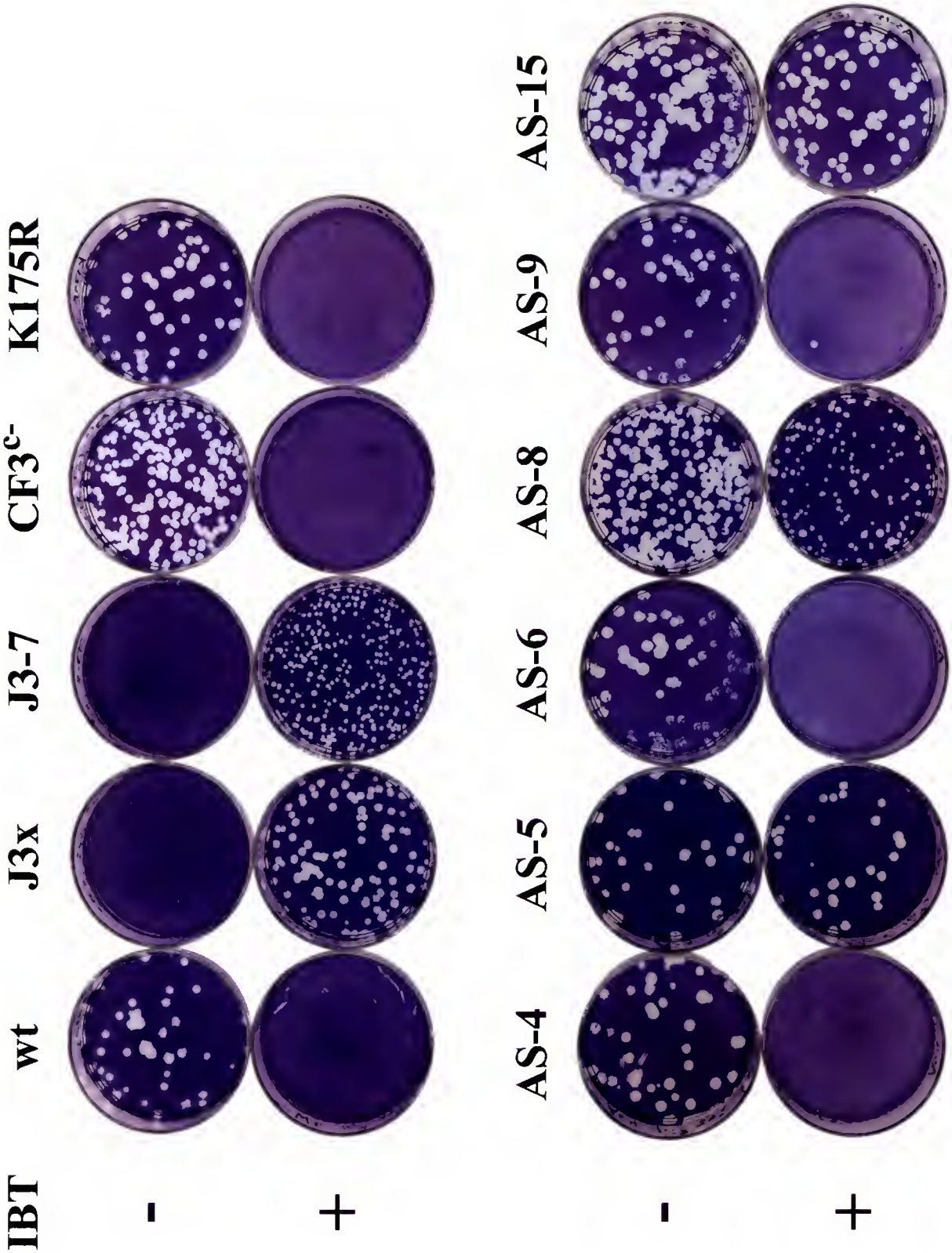
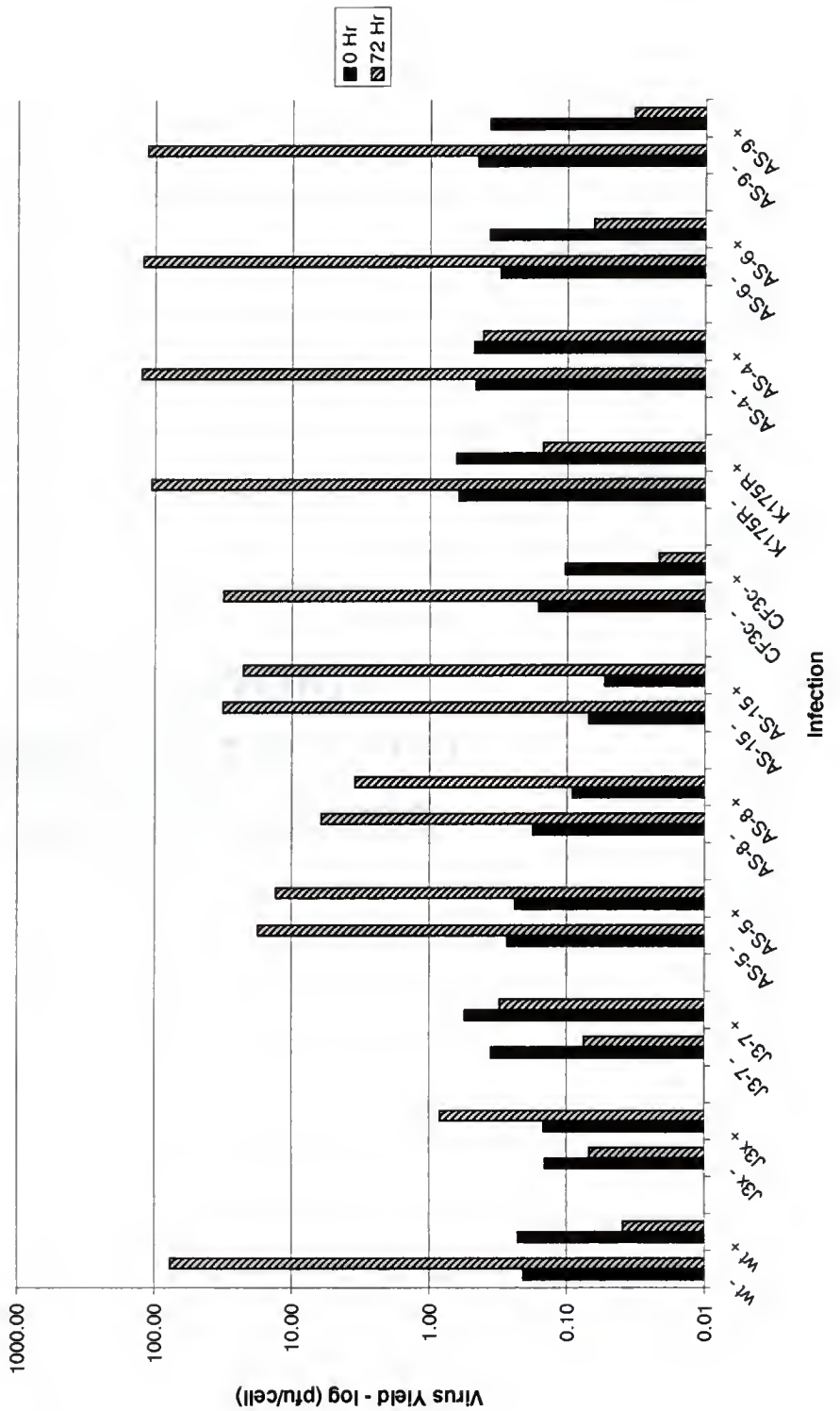

IBT
wt
J3x
J3-7
CF3c-
K175R
-
+
AS-4
AS-5
AS-6
AS-8
AS-9
AS-15
-
+

Figure 3-2. One-step growth analysis of J3 site-directed mutants. Confluent monolayers of BSC40 cells in 60-mm dishes were infected with the indicated viruses at an m.o.i. = 3.3. Infections were incubated in the presence or absence of IBT at 37°C and were harvested at 0 and 72 hrs postinfection. Lysates were then titered under permissive conditions for each virus. Specifically, all lysates with the exception of *J3x* and *J3-7* were titered at 37°C in the absence of IBT. The *J3x* and *J3-7* lysates were titered in the presence of IBT at 37°C. The numerical data presented at the left is the same data presented in the graph. Note that virus yield in the graph is shown as pfu/cell on a logarithmic scale. Solid bars indicate lysates that were harvested at 0 hrs post-infection. Striped bars represent lysates that were harvested at 72 hrs post-infection.



Infection	Yield (pfu/cell) 0 hr	72 hr
wt -	0.21	77.33
wt +	0.23	0.04
J3x -	0.14	0.07
J3x +	0.15	0.83
J3-7 -	0.35	0.07
J3-7 +	0.55	0.31
AS-5 -	0.27	17.60
AS-5 +	0.24	13.07
AS-8 -	0.18	6.13
AS-8 +	0.09	3.47
AS-15 -	0.07	32.00
AS-15 +	0.05	22.67
CF3^{c} -	0.16	31.73
CF3^{c} +	0.10	0.02
K175R -	0.61	106.67
K175R +	0.64	0.15
AS-4 -	0.46	125.33
AS-4 +	0.47	0.41
AS-6 -	0.31	122.67
AS-6 +	0.37	0.06
AS-9 -	0.45	114.67
AS-9 +	0.37	0.03

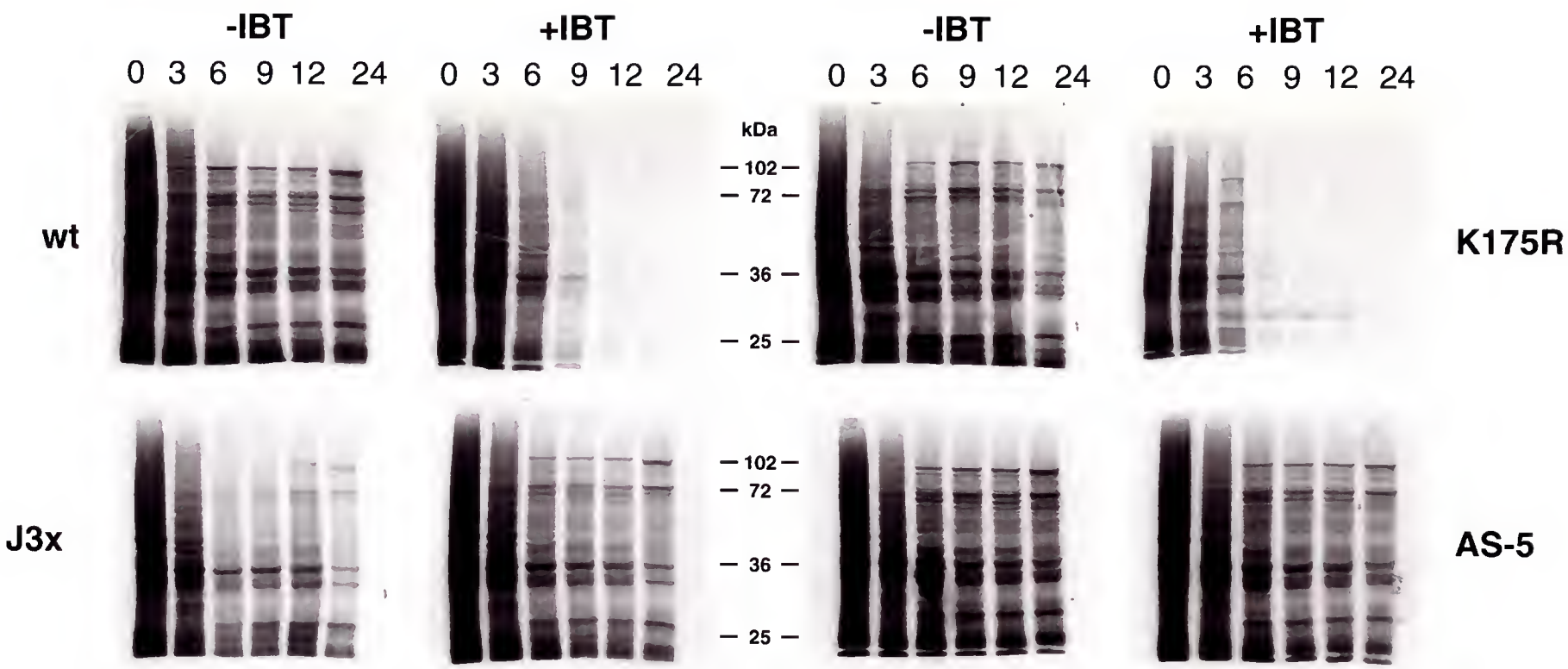


Figure 3-3. Protein synthesis in wt and mutant-infected cells. Confluent monolayers of BSC40 cells in 35-mm dishes were infected at m.o.i. = 10 with all viruses except for J3-7 which was infected at an m.o.i = 6.6 due stock titer constraints. Infected cells were incubated at 37°C in the presence or absence of IBT for 0, 3, 6, 9, 12, or 24 hrs postinfection at which time they were pulse labeled for 15 min with [35S]methionine. Infected cells were then lysed in SDS-PAGE sample buffer and solubilized proteins were and analyzed by SDS-PAGE. The gels were stained with Coomassie blue, dried, and autoradiographed. The migration of molecular size standards is shown in kilodaltons. The time postinfection (in hours) at which each sample was harvested and whether IBT was present is indicated at the top of the gels.

Figure 3-4. Northern analysis of RNA from wt and mutant infected cells. Confluent monolayers of BSC40 cells in 100-mm dishes were infected at m.o.i. = 10 with each virus except for *J3-7* which was infected at m.o.i. = 6 due to stock titer constraints. Infected cells were incubated at 37°C in the presence or absence of IBT for 3, 9, or 12 hours postinfection then total cellular RNA was purified. RNAs (1.5μg each) were electrophoresed on formaldehyde agarose gels and were transferred to nylon. The 3 hr, 9 hr, and 12 hr RNAs were respectively probed with standard early (gene C11), intermediate (gene G8), or late (gene F17) radiolabeled riboprobes. The autoradiograms are shown as follows: panel A: 3 hr RNA, early C11 probe, -IBT; panel B: 3 hr RNA, early C11 riboprobe, +IBT; panel C: 9 hr RNA, intermediate G8 probe, -IBT; panel D: 9 hr RNA, intermediate G8 probe, +IBT; panel E: 12 hr RNA, late F17 probe, -IBT; panel F: 12 hr RNA, late F17 probe, +IBT. The migration of molecular weight markers are shown at the left of each panel in kb.

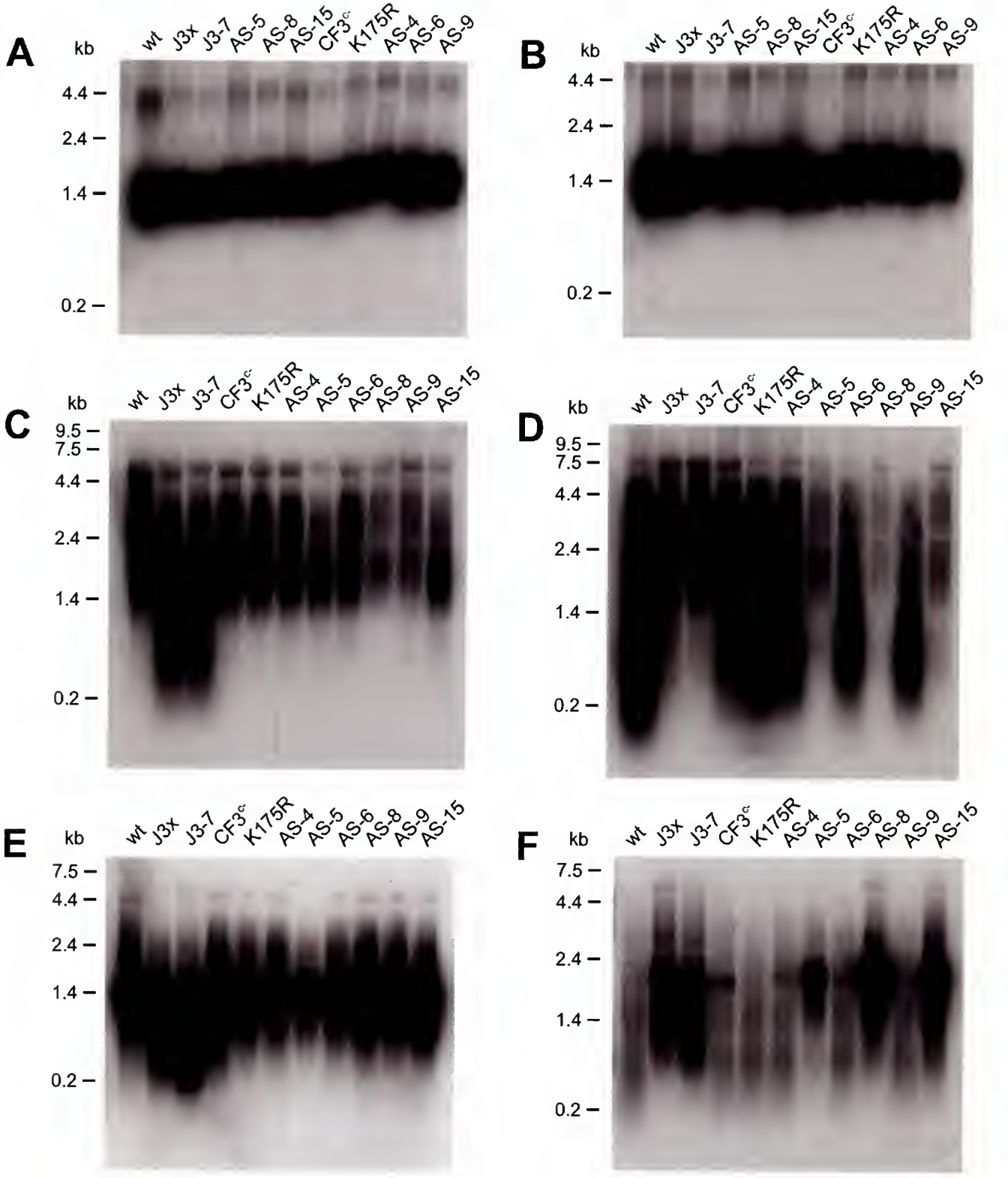

A
kb
wt
J3x
J3-7
AS-5
AS-8
AS-15
CF3c-
K175R
AS-4
AS-6
AS-9
4.4
2.4
1.4
0.2
B
kb
wt
J3x
J3-7
AS-5
AS-8
AS-15
CF3c-
K175R
AS-4
AS-6
AS-9
4.4
2.4
1.4
0.2
C
kb
wt
J3x
J3-7
CF3c-
K175R
AS-4
AS-5
AS-6
AS-8
AS-9
AS-15
9.5
7.5
4.4
2.4
1.4
0.2
D
kb
wt
J3x
J3-7
CF3c-
K175R
AS-4
AS-5
AS-6
AS-8
AS-9
AS-15
9.5
7.5
4.4
2.4
1.4
0.2
E
kb
wt
J3x
J3-7
CF3c-
K175R
AS-4
AS-5
AS-6
AS-8
AS-9
AS-15
7.5
4.4
2.4
1.4
0.2
F
kb
wt
J3x
J3-7
CF3c-
K175R
AS-4
AS-5
AS-6
AS-8
AS-9
AS-15
7.5
4.4
2.4
1.4
0.2

Figure 3-5. Analysis of poly(A) tail lengths in wt and mutant viruses. Confluent monolayers of BSC40 cells in 60-mm dishes were infected at m.o.i. = 10 with each virus except for *J3-7* which was infected at m.o.i = 6.6 due to titer constraints. Infected cells were incubated at 37°C for 12 hrs, then total cellular RNA was harvested. 1μg of total RNA was then 3' end labeled with 5μCi of [^{32}P]pCp (3000 Ci/mmol), digested with RNases A and T_1, and the labeled, RNase-resistant poly(A) tails were electrophoresed on a 10% polyacrylamide 8M urea sequencing gel. The gel was dried and subjected to autoradiography. The migration of size markers, in nucleotides, is shown at left.

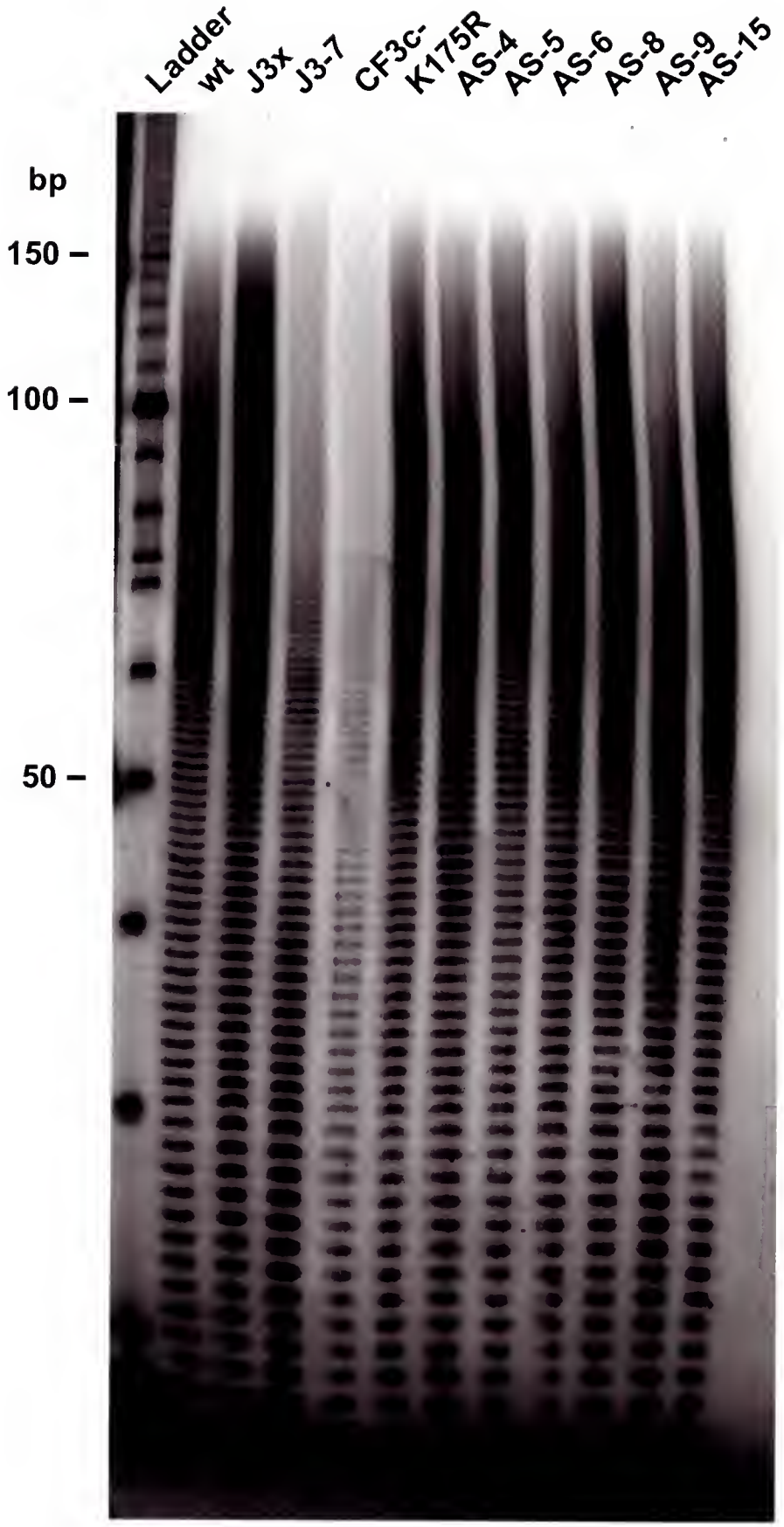

Ladder
wt
J3x
J3-7
CF3c-
K175R
AS-4
AS-5
AS-6
AS-8
AS-9
AS-15
bp
150 –
100 –
50 –

Figure 3-6. Distribution of poly(A) tail lengths in wt, *J3-7*, and CF3^{c-} samples. The distribution of poly(A) tails from the wt, *J3-7*, and CF3^{c-} samples analyzed on the gel shown in figure 5 was further examined by phosphorimager analysis. The relative amount of poly(A) tails for each sample are plotted as a function of the distance they traveled from the top of the gel. The x-axis indicates distance from the top of the gel in base pairs. The y-axis indicates relative amounts of radioactivity in arbitrary units.

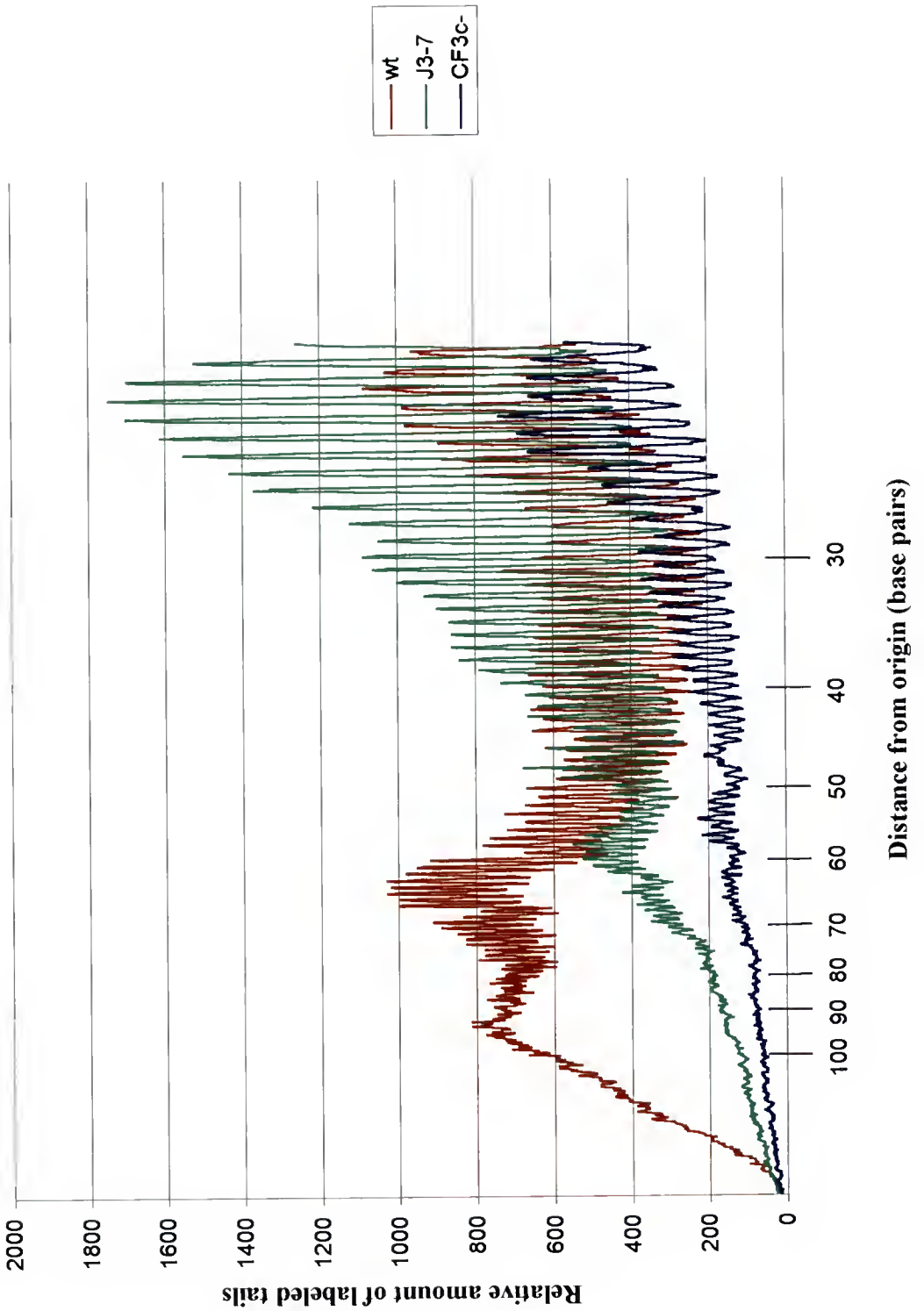

wt
J3-7
CF3c-
Relative amount of labeled tails
2000
1800
1600
1400
1200
1000
800
600
400
200
0
100 90 80 70 60 50 40 30
Distance from origin (base pairs)

Figure 3-7. *In vivo* 2'-o-methyltransferase assay. Confluent monolayers of BSC40 cells in 100-mm dishes were infected at m.o.i. = 10 with each virus except for *J3-7* which was infected at m.o.i = 6 due to titer constraints. Infections were incubated at 37°C for 10 hours, then the infected cells were washed with 6 ml of phosphate-free, serum-free 1xDME. 3.33 mCi of [^{32}P]orthophosphoric acid diluted in 4 ml of phosphate-free, serum-free 1xDME and was applied to the cells which were then incubated at 37°C for 2 additional hours. Poly(A)$^{+}$-RNA was harvested from the infected, labeled cells at 12 hours postinfection. The RNA was then digested with RNases A, T_1, T_2, and calf intestinal alkaline phosphatase. The digestion products were then electrophoresed on a 20% polyacrylamide 8M urea sequencing gel. The gel was then dried and subjected to phosphorimager analysis.

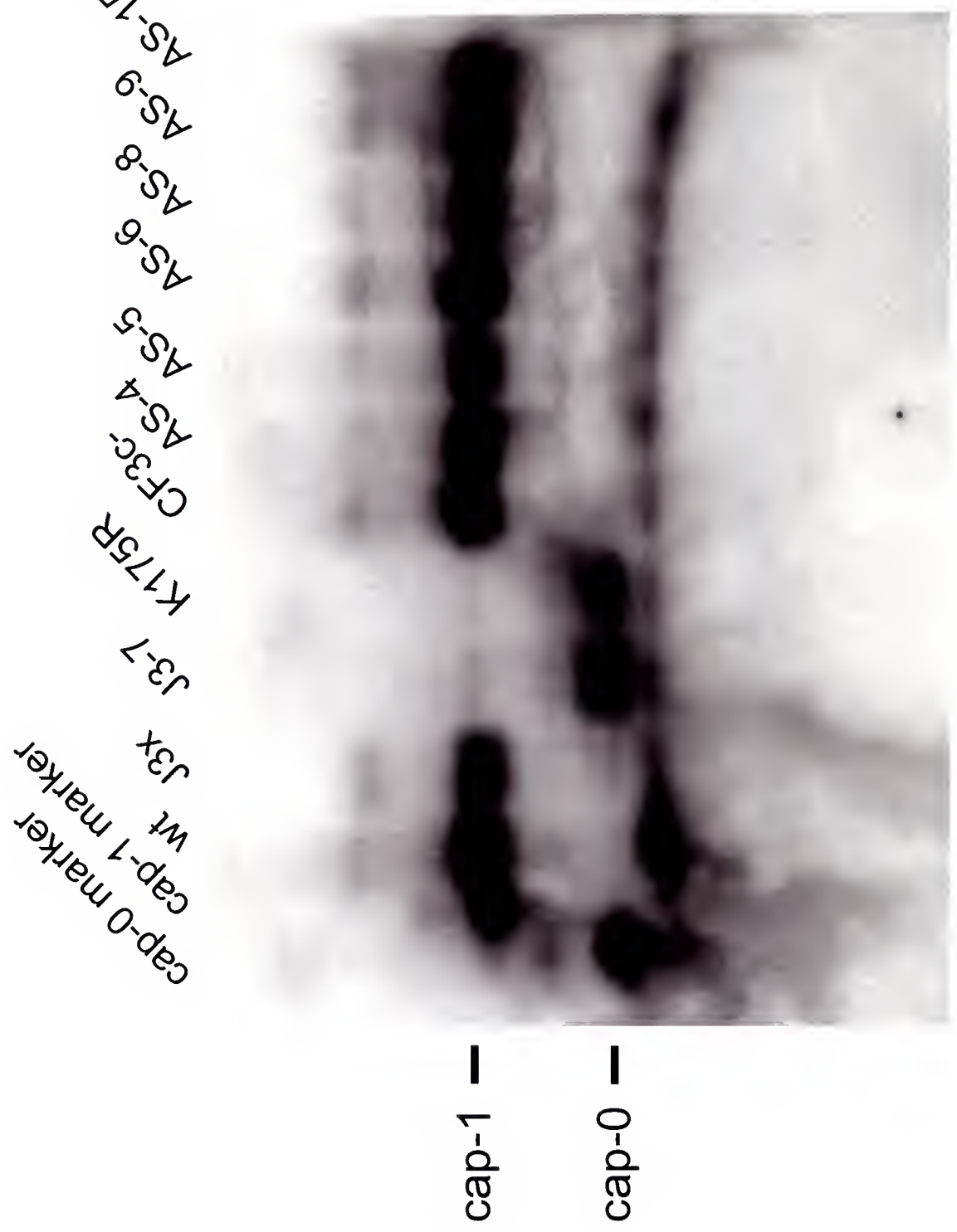
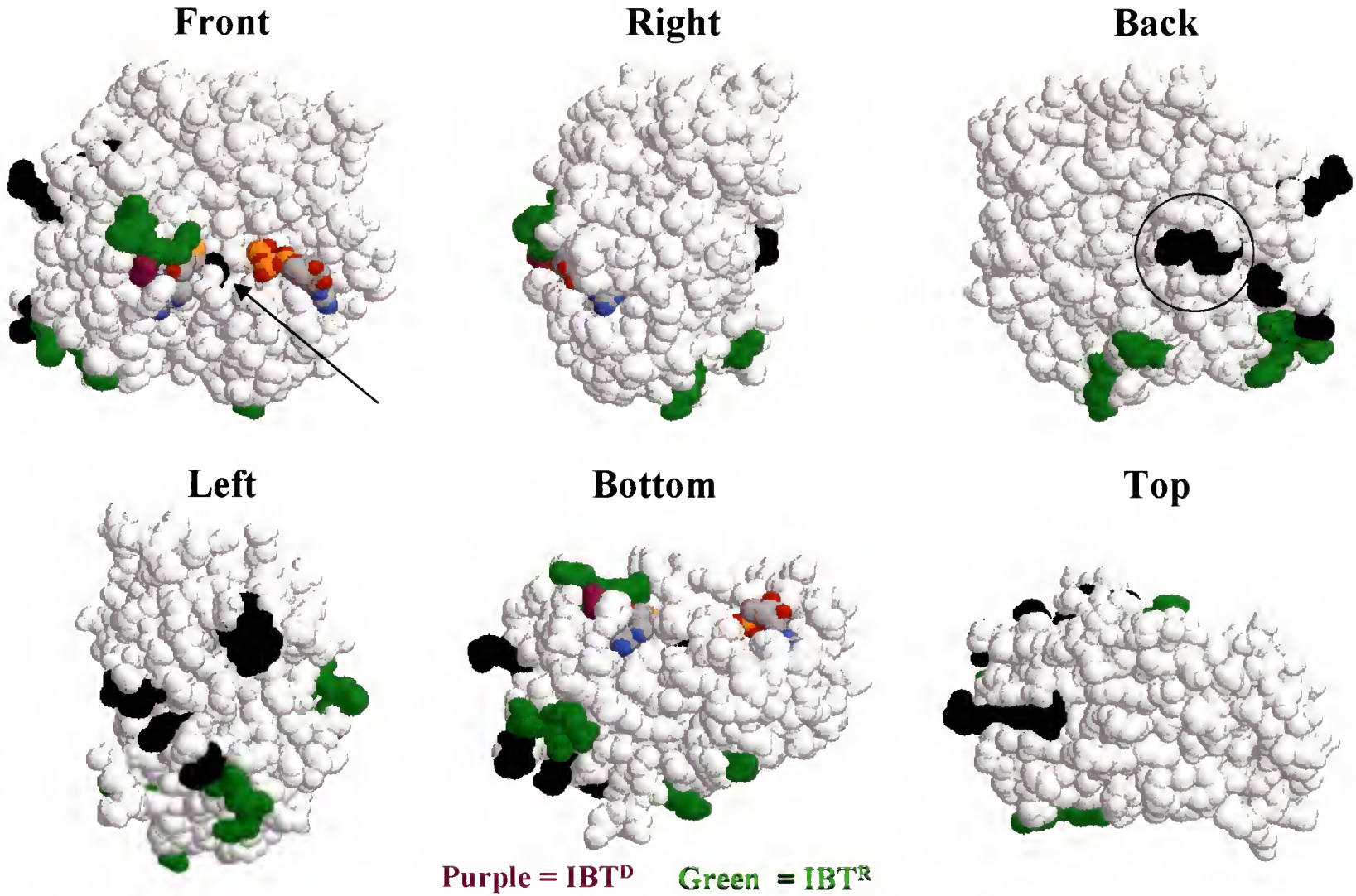

cap-0 marker
cap-1 marker
wt
J3x
J3-7
K175R
CF3c-
AS-4
AS-5
AS-6
AS-8
AS-9
AS-15
cap-1
cap-0

Figure 3-8. J3 crystal structure. The published J3 crystal structure [PDB (Protein Data Bank) identification number 1V39] is available through the National Center for Biotechnology Information at http://www.ncbi.nlm.nih.gov and is shown as a spacefilling model. The panels are labeled according to their rotational relationship to "front" view. Specifically, the right, left, bottom, and top views are rotated 90° relative to the front view. The back view is rotated 180° relative to the front. The arrow indicates the methyltransferase active site. S-adenosylhomocysetine, the product of the reaction between S-adenosylmethionine (SAM) and the 2' position of the ribose on the penultimate nucleotide, and 7-methyl G triphosphate are to the left and right of the arrow respectively. The residue highlighted in purple is glycine 96, which is mutated to aspartate in the IBT dependent *J3x* point mutant. The residues highlighted in green are the residues that are mutated to alanine in the IBT resistant mutants (AS-5, AS-8, and AS-15). The residues highlighted in black are those that have no effect on transcription when mutated to alanine (mutants AS-4, AS-6, and AS-8). The circle in the back view indicates the residues that are mutated in the $CF3^{c-}$ poly(A) polymerase stimulatory mutant. These residues are important for binding to the poly(A) polymerase. The arrow points to the K175R residue (in black) that is disrupts methyltransferase activity by inhibiting SAM binding.


Front
Right
Back
Left
Bottom
Top
Purple = IBT^D
Green = IBT^R
Black = IBT^S

CHAPTER 4
DISCUSSION

This work has identified the vaccinia virus multi-functional J3 protein as a factor that promotes the synthesis of full length viral transcripts from intermediate and late genes. Evidence for the role of J3 as a transcription factor comes from the *in vivo* biochemical characterization of mutants that were isolated through two independent genetic selections. First, a J3 point mutant, called *J3x*, was isolated as an extragenic suppressor of a temperature sensitive mutation in the A18 transcript release factor. Second, several J3 null mutations were isolated in a selection for viruses that require the transcription elongation enhancing drug IBT for growth. Examination of these J3 mutants revealed that although they have normal early gene transcription, they produce short 3' truncated postreplicative gene transcripts. The production of 3' truncated postreplicative transcripts by the J3 mutants is consistent with the absence of a positive transcription elongation factor or anti-termination factor in the mutant infections. Generalized 3' truncation of postreplicative transcripts would account for the observed synthesis of abnormally short mRNAs, and translation of abnormally short mRNAs would account for the observed specific decrease in synthesis of large but not small intermediate and late proteins (Latner et al. 2000; Xiang et al. 2000a). The phenotype is also consistent with prior genetic data: growth of the J3 mutants can be rescued by procedures that enhance transcription elongation, specifically either treatment with the anti-poxviral drug IBT or recombination with mutants in gene A18. By analogy with

previous characterization of the G2 gene, these results strongly suggest that the J3 gene is an essential positive regulator of postreplicative gene transcription elongation (Black and Condit 1996; Condit et al. 1996).

This work has also demonstrated that the J3 transcription factor activity is a third function of the protein that is independent from its two previously described activities as a 2'-o-methyltransferase and as the stimulatory subunit of the poly(A) polymerase. Evidence showing that the J3 transcription activity is a separate function of the protein comes from measurements of transcript length, poly(A) tail length, and cap methylation status of mRNAs produced *in vivo* from three different viruses. Specifically, the point mutant *J3x* was shown to be specifically defective for the stimulation transcription while the site-directed mutant K175R was shown to be specifically defective for methyltransferase activity and the site-directed mutant $CF3^{c-}$ is defective only for poly(A) polymerase stimulatory activity. In addition to showing that all three J3 functions are genetically separable, these results also demonstrate that the J3 transcription factor activity is the only J3 function that is required for virus growth, as mutations that abrogate the J3 transcription activity leave the virus dependent upon the drug IBT for growth while the K175R and $CF3^{c-}$ viruses have essentially wildtype phenotypes. Thus, neither long poly(A) tails or 2'-o-methylation of the mRNA cap is required for virus growth in tissue culture.

In attempt to determine which amino acid residues in the J3 protein are important for its transcription factor activity, several site-directed charge-to-alanine scanning mutants were constructed. Although none of the mutants that were constructed were dependent upon IBT for growth, three of those mutants, called AS-5, AS-8, and AS-15,

were resistant to IBT. Noting that the IBT resistant growth phenotype is an intermediate between the dependent and sensitive phenotypes, it was initially hypothesized that the IBT resistant mutants would synthesize postreplicative gene transcripts that are longer than those produced by the IBT dependent viruses, but shorter than the IBT sensitive wildtype virus. Analysis of the transcripts produced by the IBT resistant viruses in the absence of drug revealed that they are not obviously different from those produced by the wildtype virus, suggesting that any transcription defects they may have in the absence of drug are subtle and may require a more sensitive assay for better characterization. In the presence of drug, the IBT sensitive viruses experienced RNA breakdown which has previously been ascribed to induction of the 2-5A RNA degradation pathway. This pathway is turned on by the accumulation of double-stranded RNA resulting from the over-stimulation of transcription elongation by IBT. In contrast to the IBT sensitive viruses, the IBT resistant mutants did not experience significant RNA degradation in the presence of drug, but synthesized transcripts that were of a size and quantity that was similar to the transcripts they produced in the absence of drug. Moreover, these transcripts were able to support normal translation. In the presence of IBT the phenotype of the resistant viruses was similar to that of the IBT dependent mutants. These observations suggest that the IBT resistant viruses do in fact have some sort of alteration in the transcription apparatus that is similar to the alteration in the dependent viruses, which include those containing null mutation of the J3 gene.

The construction and analysis of the site-directed mutants provided additional insights into the role that J3 has in transcription. First, it appears that the J3 transcription factor activity may not require that J3 form a heterodimer with the E1 poly(A)

polymerase. Evidence for this comes from previous work which has suggested that that the defect in poly(A) tail synthesis displayed by the $CF3^{c-}$ protein in stimulating the E1 poly(A) polymerase elongation *in vitro* is the result of the protein's inability to bind E1 (Shi et al. 1997). Assuming that the $CF3^{c-}$ mutant J3 protein does not bind E1 *in vivo* and knowing that it has normal transcription activity suggests that J3 does not require heterodimer formation with E1 for stimulating RNA polymerase transcription. Second, a surface area of J3 that might be important for mediating its effects on transcription has been highlighted by mapping the mutations in the IBT dependent *J3x* virus and the IBT resistant AS-5, AS-8, and AS-15 viruses on the crystal structure. As shown in figure 3-8, the purple and green residues indicate mutations that lead to an altered response to IBT and are all located in roughly the same hemisphere on the molecule. This could be an area of the protein that is important for binding an additional transcription factor or the polymerase. In fact, work by others has been shown that J3 binds to three other transcription factors besides the poly(A) polymerase. Specifically, J3 binds to NPH-I, H4, and A18 (Mohamed et al. 2001; Ed Niles, personal communication). NPH-I (nucleoside triphosphate phosphohydrolase I) is an ATPase that is required in addition to the heterodimeric capping enzyme and H4 for termination of early gene transcription and transcript release (Mohamed and Niles 2000; Mohamed and Niles 2001). H4 (RAP94) is a polymerase subunit that is specific for early genes. H4 is required for early gene transcription initiation and for response to the early gene termination signal U_5NU (Mohamed and Niles 2001). Although J3 is not required for early gene transcription elongation or termination, its association with NPH-I and H4 may position it on the polymerase to facilitate its activities in cap methylation and poly(A) addition, thus

linking transcription to mRNA processing. NPH-I and H4 likewise have no known role in post-replicative gene transcription. Interestingly, GST-pulldown experiments were used to roughly map the regions of J3 that bind with NPH-I and H4. Specifically, NPH-I binds J3 at amino acids 169-249 and H4 binds J3 at amino acids 169-333 (Mohamed et al. 2001). These regions are highlighted on the J3 crystal structure shown in figure 4-1. While there is some overlap between the areas that are important for binding NPH-I and H4 with two residues that are mutated in the resistant mutant AS-15, it appears that most of the residues that are mutated in the IBT resistant viruses are outside of the NPH-I and H4 binding domains. This provides support for the hypothesis that NPH-I and H4 simply provide a docking site for J3, linking it to the transcription complex and positioning it for its 2'-o-methyltransferase and poly(A) polymerase stimulatory activities. It also supports previous data suggesting that NPH-I and H4 do not have a role in postreplicative gene transcription. In addition, this observation leaves open the possibility that J3 could bind an additional factor that is important for mediating its transcription activity on a surface area of the protein that is distinct from its NPH-I and H4 binding sites. As mentioned, some preliminary evidence suggests that J3 can also bind to the A18 protein (Ed Niles, personal communication). A18 is a DNA-dependent ATPase and 3'-5' DNA helicase that is required for release of transcripts that are initiated from all three classes of promoters (Lackner and Condit 2000). Interestingly, the G2 transcription factor also binds A18 (Black et al. 1998). The exact nature of the interactions between J3, G2, and A18 awaits further characterization.

Besides J3, G2 and A18, there are likely to be additional factors that are involved in the regulation of intermediate and late gene transcription elongation and termination.

This prediction is based on experiments with the drug IBT showing that resistance to IBT can arise from mutation of A24 which is the second largest subunit of the polymerase (Condit et al. 1991) and from mutation of another unmapped gene which is not A24, J3, G2, or A18 (Steve Cresawn, personal communication). Neither the target of IBT or its mechanism of action are known. However, its structure (Fig. 4-2) resembles a purine with a side chain attached to the five-member ring. In a previously published study (Katz 1987), it was shown that the $-NH_2$ group on the side chain was required for inhibition of wildtype virus growth and for supporting growth of an unmapped dependent mutant. It is likely that G2 and J3 are not the direct targets of IBT because null mutations in either protein cause the virus to become dependent upon the drug for growth. It has also been shown that IBT has no effect when directly added to an *in vitro* transcription assay that was designed to measure A18 release factor activity (Cari Lackner, personal communication). Thus, a metabolized form of IBT might be responsible for the drug's effects on transcription.

In spite of the data, a definition for the mechanistic role of J3 in stimulating postreplicative gene transcription remains unclear. The *E.coli* and eukaryotic post-initiation transcription factors described in chapter 1 generally fall into one of two broad categories: 1) those that enhance elongation and 2) those that induce or inhibit termination. Currently, there is not enough information to indicate if J3 is acting at the elongation or termination stage. The transcription defects observed in the J3 null or IBT dependent mutants could be explained if the normal function of J3 is to stimulate the rate of elongation or to suppress pausing either through modification of the polymerase structure or by inducing the intrinsic transcript cleavage activity of the polymerase.

Alternatively, it could prevent termination through direct interaction with release factors like A18.

The fact that J3 stimulates the rate of E1 poly(A) polymerase elongation suggests that it might also stimulate the elongation rate of the RNA polymerase. However, given the mechanism by which it is thought to stimulate E1, it is not likely to function in the exact same capacity for the RNA polymerase. From several published experiments, it appears that *in vitro*, E1 binds multiple U residues located within 31-40 nucleotides of the 3' end of non-polyadenylated transcripts (Deng et al. 1997; Gershon 1998). By itself, E1 will add about 35 A residues to the 3' end before it essentially stops. Some evidence suggests that the reason E1 stops elongation is that in the process of adding A residues, it separates itself from the U residues that it needs for binding to the transcript. When J3 binds to E1, a RNA binding groove or channel is formed in the cleft along the dimerization interface between the two proteins and thus helps E1 stay associated with the poly(A) tail (Gershon 1998). Thus, as an elongation factor, it is difficult to imagine J3 performing the same function for the RNA polymerase as it does for E1. However, it is possible that J3, like the pol II elongation factors elongin, ELL, and CSB, may stimulate the rate of elongation through an alternate mechanism that has not yet been defined. So far, the only reason to suspect that it may specifically act to prevent termination is preliminary evidence showing that it binds to the A18 release factor (Ed Niles, personal communication).

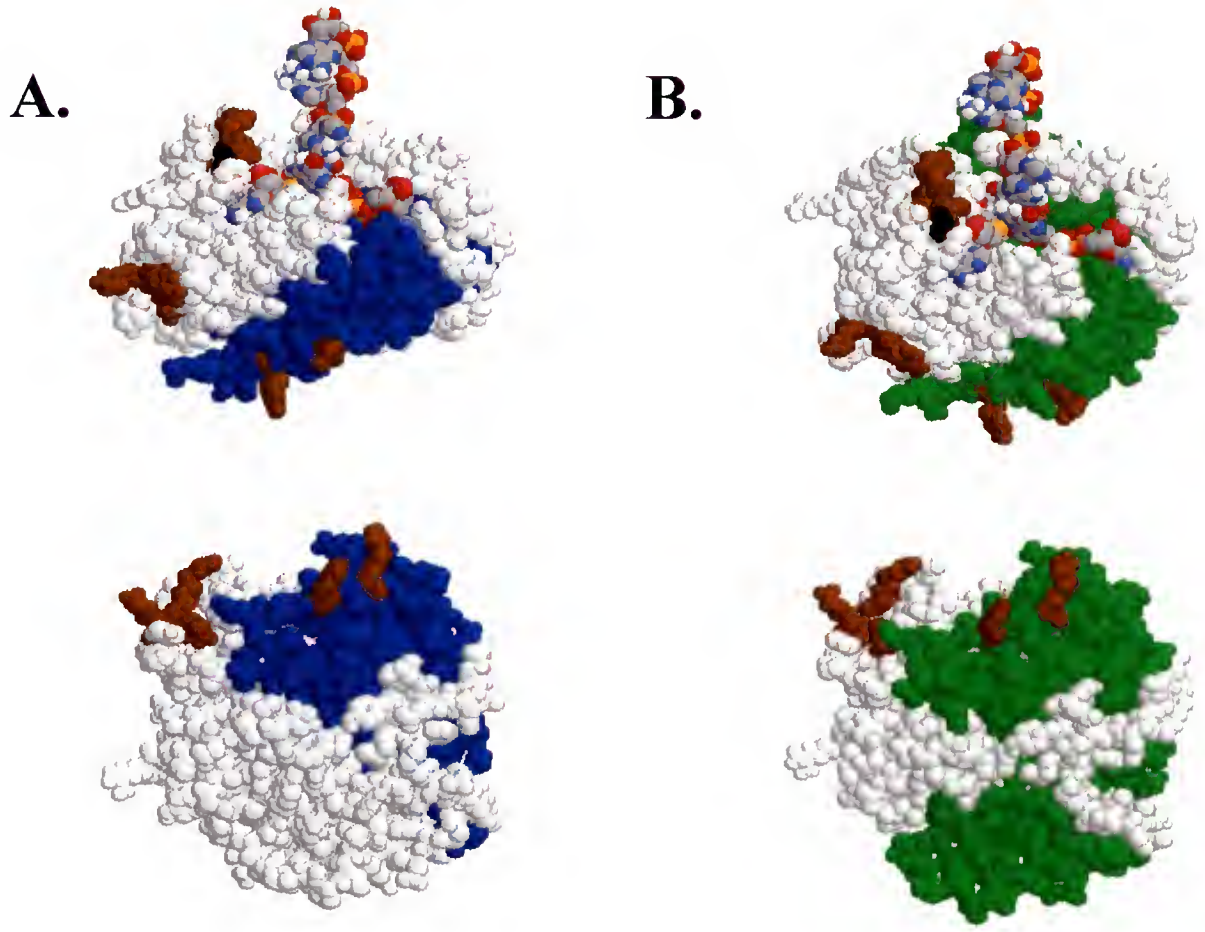
Establishing the precise mechanism by which J3 enhances transcription will probably require an *in vitro* assay to measure its activity at each stage of transcript synthesis. Such an assay might be similar to that used to characterize the transcript

release activity of the A18 protein (Lackner and Condit 2000). Typically, these kinds of assays consist of forming transcription complexes on magnetic-bead-bound DNA templates using infected cell extracts. Extracts from J3 mutant infected cells or extracts that have been immunodepeleted of J3 could be used to initiate transcription. The polymerase complex can then be manipulated in a number of ways as it traverses the template. For example, it can be paused by varying the NTP concentrations or the template sequence. The complexes can then be washed in high salt buffer to strip off loosely associated proteins. The transcripts and polymerase complexes that remain associated with the template can recovered on a magnet while the terminated (or released) transcripts can be recovered in the supernatant. Subsequent addition of various purified recombinant proteins, such as J3, G2, A18, or partially purified cell fractions can be added to examine the effects on transcription. For example, if J3 is an anti-termination factor that directly inhibits A18 mediated release, it might be hypothesized that the A18 release activity could be titrated by the addition of excess J3 protein. In fact, some preliminary experiments of this type have been performed but have so far been uninformative (Cari Lackner and Steve Cresawn, personal communication). It remains to be determined if variations on the reaction conditions, such as salt and nucleotide concentrations, template sequence variations, or freshly prepared recombinant protein might yield more informative results. Alternatively, an analysis of the transcript lengths produced as a function of time could be used to determine if J3 affects the rate of elongation. Other types of experiments that could help define the role of J3 in transcription might include a better characterization of the protein-protein interactions that J3 has with other factors like A18. This could be done by GST-pulldown

experiments using recombinant J3 and A18 truncation or deletion mutant proteins or by yeast two-hybrid analysis. Finally, the mapping of additional mutations that generate IBT resistance or dependence might lead to the identification of other important transcription factors.

Figure 4-1. NPH-1 and H4 interaction domains. The published J3 crystal structure [PDB (Protein Data Bank) identification number 1AV6] is available through the National Center for Biotechnology Information at http://www.ncbi.nlm.nih.gov and is shown as a spacefilling model. Amino acid residues 169- 249 are required for binding to NPH-I according to GST-pulldown experiments and are highlighted in blue in panel A. Amino acid residues 169-333 are required for binding to H4 according to GST-pulldown experiments and are highlighted in green in panel B. In both panels, the structure on the bottom is rotated 90° backwards on the horizontal axis relative to the top. Also, in both panels, the residue that is highlighted in black is the glycine 96, which is mutated to an aspartate in the IBT dependent mutant *J3x*. The residues shown in brown are mutated to alanine in the IBT resistant mutants. AS-5, AS-8, and AS-15. A 7-methyl-G-triphosphate capped RNA hexamer is shown in the active site of the methyltransferase domain with S-adenosylhomocysteine, the product of the reaction between the methyl donor S-adenosylmethionine and the 2' position of the first transcribed nucleotide.

NPH-I and H4 Interaction Domains



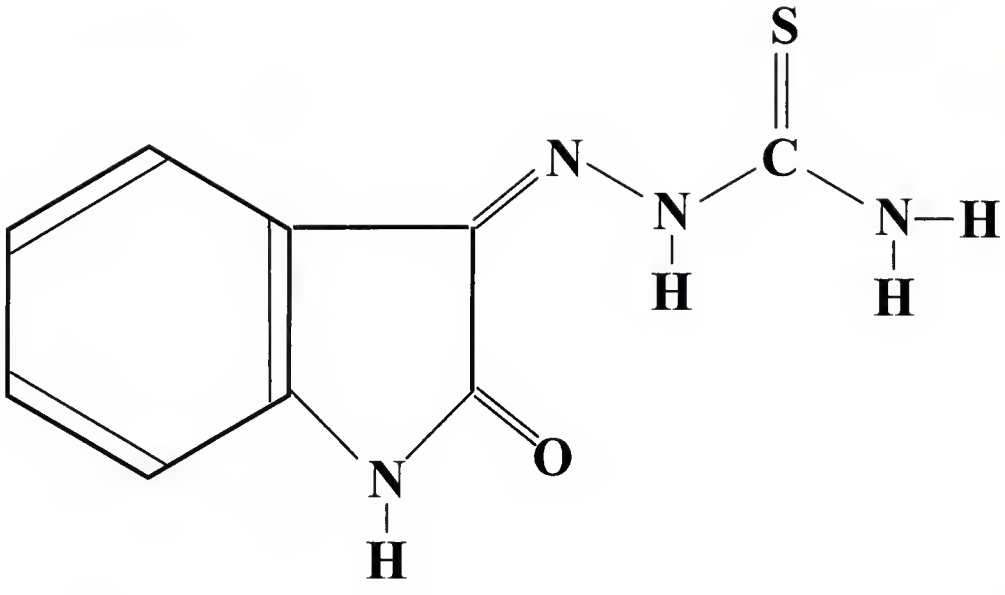


Figure 4-2. Molecular structure of isatin-β-thiosemicarbazone (IBT). The sulfur atom and the amine group in the side chain are largely responsible for inhibition of wt virus growth and for supporting the growth of viruses that are dependent upon IBT.

LIST OF REFERENCES

Afonso,C.L., E.R.Tulman, Z.Lu, E.Oma, G.F.Kutish, and D.L.Rock. 1999. The genome of Melanoplus sanguinipes entomopoxvirus. J.Virol. **73**: 533-552.

Ahn,B.Y., P.D.Gershon, and B.Moss. 1994. RNA polymerase-associated protein Rap94 confers promoter specificity for initiating transcription of vaccinia virus early stage genes. J.Biol.Chem. **269**: 7552-7557.

Ahn,B.Y., and B.Moss. 1989. Capped poly(A) leaders of variable lengths at the 5' ends of vaccinia virus late mRNAs. J.Virol. **63**: 226-232.

Aranda,A., and N.Proudfoot. 2001. Transcriptional termination factors for RNA polymerase II in yeast. Mol.Cell **7**: 1003-1011.

Ashfield,R., P.Enriquez-Harris, and N.J.Proudfoot. 1991. Transcriptional termination between the closely linked human complement genes C2 and factor B: common termination factor for C2 and c-myc? EMBO J. **10**: 4197-4207.

Ashfield,R., A.J.Patel, S.A.Bossone, H.Brown, R.D.Campbell, K.B.Marcu, and N.J.Proudfoot. 1994. MAZ-dependent termination between closely spaced human complement genes. EMBO J. **13**: 5656-5667.

Baldick,C.J., Jr., J.G.Keck, and B.Moss. 1992. Mutational analysis of the core, spacer, and initiator regions of vaccinia virus intermediate-class promoters. J.Virol. **66**: 4710-4719.

Barbosa,E., and B.Moss. 1978a. mRNA(nucleoside-2'-)-methyltransferase from vaccinia virus. Characteristics and substrate specificity. J.Biol.Chem. **253**: 7698-7702.

Barbosa,E., and B.Moss. 1978b. mRNA(nucleoside-2'-)-methyltransferase from vaccinia virus. Purification and physical properties. J.Biol.Chem. **253**: 7692-7697.

Baroudy,B.M., and B.Moss. 1980. Purification and characterization of a DNA-dependent RNA polymerase from vaccinia virions. J.Biol.Chem. **255**: 4372-4380.

Bawden,A.L., K.J.Glassberg, J.Diggans, R.Shaw, W.Farmerie, and R.W.Moyer. 2000. Complete genomic sequence of the Amsacta moorei entomopoxvirus: analysis and comparison with other poxviruses. Virology **274**: 120-139.

Bayliss,C.D., and R.C.Condit. 1993. Temperature-sensitive mutants in the vaccinia virus A18R gene increase double-stranded RNA synthesis as a result of aberrant viral transcription. Virology **194**: 254-262.

Bayliss,C.D., and R.C.Condit. 1995. The vaccinia virus A18R gene product is a DNA-dependent ATPase. J.Biol.Chem. **270**: 1550-1556.

Beaud,G., and R.Beaud. 1997. Preferential virosomal location of underphosphorylated H5R protein synthesized in vaccinia virus-infected cells. J.Gen.Virol. **78 (Pt 12)**: 3297-3302.

Birse,C.E., L.Minvielle-Sebastia, B.A.Lee, W.Keller, and N.J.Proudfoot. 1998a. Coupling termination of transcription to messenger RNA maturation in yeast. Science **280**: 298-301.

Birse,C.E., L.Minvielle-Sebastia, B.A.Lee, W.Keller, and N.J.Proudfoot. 1998b. Coupling termination of transcription to messenger RNA maturation in yeast. Science **280**: 298-301.

Black,E.P., and R.C.Condit. 1996. Phenotypic characterization of mutants in vaccinia virus gene G2R, a putative transcription elongation factor. J.Virol. **70**: 47-54.

Black,E.P., N.Moussatche, and R.C.Condit. 1998. Characterization of the interactions among vaccinia virus transcription factors G2R, A18R, and H5R. Virology **245**: 313-322.

Broyles,S.S., and J.Li. 1993. The small subunit of the vaccinia virus early transcription factor contacts the transcription promoter DNA. J.Virol. **67**: 5677-5680.

Broyles,S.S., J.Li, and B.Moss. 1991. Promoter DNA contacts made by the vaccinia virus early transcription factor. J.Biol.Chem. **266**: 15539-15544.

Broyles,S.S., and B.Moss. 1987. Sedimentation of an RNA polymerase complex from vaccinia virus that specifically initiates and terminates transcription. Mol.Cell Biol. **7**: 7-14.

Broyles,S.S., L.Yuen, S.Shuman, and B.Moss. 1988. Purification of a factor required for transcription of vaccinia virus early genes. J.Biol.Chem. **263**: 10754-10760.

Burgess,B.R., and J.P.Richardson. 2001. RNA passes through the hole of the protein hexamer in the complex with the Escherichia coli Rho factor. J.Biol.Chem. **276**: 4182-4189.

Cameron,C., S.Hota-Mitchell, L.Chen, J.Barrett, J.X.Cao, C.Macaulay, D.Willer, D.Evans, and G.McFadden. 1999. The complete DNA sequence of myxoma virus. Virology **264**: 298-318.

Chedin,S., M.Riva, P.Schultz, A.Sentenac, and C.Carles. 1998. The RNA cleavage activity of RNA polymerase III is mediated by an essential TFIIS-like subunit and is important for transcription termination. Genes Dev. **12**: 3857-3871.

Chodchoy,N., N.B.Pandey, and W.F.Marzluff. 1991. An intact histone 3'-processing site is required for transcription termination in a mouse histone H2a gene. Mol.Cell Biol. **11**: 497-509.

Christen,L.M., M.Sanders, C.Wiler, and E.G.Niles. 1998. Vaccinia virus nucleoside triphosphate phosphohydrolase I is an essential viral early gene transcription termination factor. Virology **245**: 360-371.

Cohrs,R.J., R.C.Condit, R.F.Pacha, C.L.Thompson, and O.K.Sharma. 1989. Modulation of ppp(A2'p)nA-dependent RNase by a temperature-sensitive mutant of vaccinia virus. J.Virol. **63**: 948-951.

Conaway,J.W., A.Dvir, R.J.Moreland, Q.Yan, B.J.Elmendorf, S.Tan, and R.C.Conaway. 1998b. Mechanism of promoter escape by RNA polymerase II. Cold Spring Harb.Symp.Quant.Biol. **63**: 357-364.

Conaway,J.W., A.Dvir, R.J.Moreland, Q.Yan, B.J.Elmendorf, S.Tan, and R.C.Conaway. 1998a. Mechanism of promoter escape by RNA polymerase II. Cold Spring Harb.Symp.Quant.Biol. **63**: 357-364.

Conaway,J.W., A.Shilatifard, A.Dvir, and R.C.Conaway. 2000a. Control of elongation by RNA polymerase II. Trends Biochem.Sci. **25** : 375-380.

Conaway,J.W., A.Shilatifard, A.Dvir, and R.C.Conaway. 2000b. Control of elongation by RNA polymerase II. Trends Biochem.Sci. **25** : 375-380.

Conaway,R.C., and J.W.Conaway. *Transcription :mechanisms and regulation*, Raven Press, New York.

Condit,R.C., R.Easterly, R.F.Pacha, Z.Fathi, and R.J.Meis. 1991. A vaccinia virus isatin-beta-thiosemicarbazone resistance mutation maps in the viral gene encoding the 132-kDa subunit of RNA polymerase. Virology **185**: 857-861.

Condit,R.C. and A.Motyczka. 1981. Isolation and preliminary characterization of temperature-sensitive mutants of vaccinia virus. Virology **113**: 224-241.

Condit,R.C., A.Motyczka, and G.Spizz. 1983. Isolation, characterization, and physical mapping of temperature- sensitive mutants of vaccinia virus. Virology **128**: 429-443.

Condit,R.C., Y.Xiang, and J.I.Lewis. 1996. Mutation of vaccinia virus gene G2R causes suppression of gene A18R ts mutants: implications for control of transcription. Virology **220**: 10-19.

Cooper,J.A., B.Moss, and E.Katz. 1979. Inhibition of vaccinia virus late protein synthesis by isatin-beta- thiosemicarbazone: characterization and in vitro translation of viral mRNA. Virology **96**: 381-392.

Cooper,J.A., R.Wittek, and B.Moss. 1981. Extension of the transcriptional and translational map of the left end of the vaccinia virus genome to 21 kilobase pairs. J.Virol. **39**: 733-745.

Cotten,M., O.Gick, A.Vasserot, G.Schaffner, and M.L.Birnstiel. 1988. Specific contacts between mammalian U7 snRNA and histone precursor RNA are indispensable for the in vitro 3' RNA processing reaction. EMBO J. **7**: 801-808.

Cuello,P., D.C.Boyd, M.J.Dye, N.J.Proudfoot, and S.Murphy. 1999. Transcription of the human U2 snRNA genes continues beyond the 3' box in vivo. EMBO J. **18**: 2867-2877.

Dantonel,J.C., K.G.Murthy, J.L.Manley, and L.Tora. 1997. Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. Nature **389**: 399-402.

Davison,A.J., and B.Moss. 1989b. Structure of vaccinia virus early promoters. J.Mol.Biol. **210**: 749-769.

Davison,A.J., and B.Moss. 1989a. Structure of vaccinia virus late promoters. J.Mol.Biol. **210**: 771-784.

de Vegvar,H.E., E.Lund, and J.E.Dahlberg. 1986. 3' end formation of U1 snRNA precursors is coupled to transcription from snRNA promoters. Cell **47**: 259-266.

DeMasi,J., and P.Traktman. 2000. Clustered charge-to-alanine mutagenesis of the vaccinia virus H5 gene: isolation of a dominant, temperature-sensitive mutant with a profound defect in morphogenesis. J.Virol. **74**: 2393-2405.

Deng,L., L.Beigelman, J.Matulic-Adamic, A.Karpeisky, and P.D.Gershon. 1997. Specific recognition of an rU2-N15-rU motif by VP55, the vaccinia virus poly(A) polymerase catalytic subunit. J.Biol.Chem. **272**: 31542-31552.

Deng,L., and S.Shuman. 1994. A role for the H4 subunit of vaccinia RNA polymerase in transcription initiation at a viral early promoter. J.Biol.Chem. **269**: 14323-14328.

Deng,L., and S.Shuman. 1998. Vaccinia NPH-I, a DExH-box ATPase, is the energy coupling factor for mRNA transcription termination. Genes Dev. **12**: 538-546.

Ding,H.F., M.Bustin, and U.Hansen. 1997. Alleviation of histone H1-mediated transcriptional repression and chromatin compaction by the acidic activation region in chromosomal protein HMG-14. Mol.Cell Biol. **17**: 5843-5855.

Ding,H.F., S.Rimsky, S.C.Batson, M.Bustin, and U.Hansen. 1994. Stimulation of RNA polymerase II elongation by chromosomal protein HMG- 14. Science **265**: 796-799.

Dvir,A., J.W.Conaway, and R.C.Conaway. 2001. Mechanism of transcription initiation and promoter escape by RNA polymerase II. Curr.Opin.Genet.Dev. **11**: 209-214.

Dvir,A., R.C.Conaway, and J.W.Conaway. 1996. Promoter escape by RNA polymerase II. A role for an ATP cofactor in suppression of arrest by polymerase at promoter-proximal sites. J.Biol.Chem. **271**: 23352-23356.

Dye,M.J., and N.J.Proudfoot. 1999. Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. Mol.Cell **3**: 371-378.

Dye,M.J., and N.J.Proudfoot. 2001. Multiple transcript cleavage precedes polymerase release in termination by RNA polymerase II. Cell **105**: 669-681.

Eisen,A., and J.C.Lucchesi. 1998. Unraveling the role of helicases in transcription. Bioessays **20**: 634-641.

Enriquez-Harris,P., N.Levitt, D.Briggs, and N.J.Proudfoot. 1991. A pause site for RNA polymerase II is associated with termination of transcription. EMBO J. **10**: 1833-1842.

Ensinger,M.J., J.P.Weir, and B.Moss. 1985. Fine structure marker rescue of temperature-sensitive mutations of vaccinia virus within a central conserved region of the genome. J.Virol. **56**: 1027-1029.

Esposito,J., R.Condit, and J.Obijeski. 1981. The preparation of orthopoxvirus DNA. J.Virol.Methods **2**: 175-179.

Evers,R., A.Smid, U.Rudloff, F.Lottspeich, and I.Grummt. 1995. Different domains of the murine RNA polymerase I-specific termination factor mTTF-I serve distinct functions in transcription termination. EMBO J. **14**: 1248-1256.

Falkner,F.G., and B.Moss. 1990. Transient dominant selection of recombinant vaccinia viruses. J.Virol. **64**: 3108-3111.

Fields,B.N., D.M.Knipe, and P.M.Howley. *Fundamental virology*, Lippincott-Raven, Philadelphia.

Fong,N., and D.L.Bentley. 2001. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. Genes Dev. **15**: 1783-1795.

Gallie,D.R. 1998. A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. Gene **216**: 1-11.

Geiduschek,E.P., and G.A.Kassavetis. 2001. The RNA polymerase III transcription apparatus. J.Mol.Biol. **310**: 1-26.

Geiselmann,J., Y.Wang, S.E.Seifried, and P.H.von Hippel. 1993. A physical model for the translocation and helicase activities of Escherichia coli transcription termination protein Rho. Proc.Natl.Acad.Sci.U.S.A **90**: 7754-7758.

Gershon,P.D. 1998. mRNA 3' End Formation by Vaccinia Virus: Mechanism of Action of a Heterodimeric Poly(A) Polymerase. Seminars in Virology **8**: 343-350.

Gershon,P.D., and B.Moss. 1990. Early transcription factor subunits are encoded by vaccinia virus late genes. Proc.Natl.Acad.Sci.U.S.A **87**: 4401-4405.

Gershon,P.D., and B.Moss. 1992. Transition from rapid processive to slow nonprocessive polyadenylation by vaccinia virus poly(A) polymerase catalytic subunit is regulated by the net length of the poly(A) tail. Genes Dev. **6**: 1575-1586.

Gershon,P.D., X.Shi, and A.E.Hodel. 1998. Evidence that the RNA methylation and poly(A) polymerase stimulatory activities of vaccinia virus protein VP39 do not impinge upon one another. Virology **246**: 253-265.

Gick,O., A.Kramer, W.Keller, and M.L.Birnstiel. 1986. Generation of histone mRNA 3' ends by endonucleolytic cleavage of the pre-mRNA in a snRNP-dependent in vitro reaction. EMBO J. **5**: 1319-1326.

Graham,J.E., and J.P.Richardson. 1998. rut Sites in the nascent transcript mediate Rho-dependent transcription termination in vivo. J.Biol.Chem. **273**: 20764-20769.

Greenblatt,J., J.R.Nodwell, and S.W.Mason. 1993. Transcriptional antitermination. Nature **364**: 401-406.

Gu,X., and W.F.Marzluff. 1996. 3' Processing and termination of mouse histone transcripts synthesized in vitro by RNA polymerase II. Nucleic Acids Res. **24**: 3797-3805.

Guan,K.L., and J.E.Dixon. 1991. Eukaryotic proteins expressed in Escherichia coli: an improved thrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase. Anal.Biochem. **192** : 262-267.

Gunasinghe,S.K., A.E.Hubbs, and C.F.Wright. 1998. A vaccinia virus late transcription factor with biochemical and molecular identity to a human cellular protein. J.Biol.Chem. **273**: 27524-27530.

Gunnery,S., Y.Ma, and M.B.Mathews. 1999. Termination sequence requirements vary among genes transcribed by RNA polymerase III. J.Mol.Biol. **286**: 745-757.

Hagler,J., and S.Shuman. 1993. Nascent RNA cleavage by purified ternary complexes of vaccinia RNA polymerase. J.Biol.Chem. **268**: 2166-2173.

Hampsey,M. 1998. Molecular genetics of the RNA polymerase II general transcriptional machinery. Microbiol.Mol.Biol.Rev. **62**: 465-503.

Hassett,D.E., and R.C.Condit. 1994. Targeted construction of temperature-sensitive mutations in vaccinia virus by replacing clustered charged residues with alanine. Proc.Natl.Acad.Sci.U.S.A **91**: 4554-4558.

Henkin,T.M. 1996. Control of transcription termination in prokaryotes. Annu.Rev.Genet. **30**: 35-57.

Henkin,T.M. 2000. Transcription termination control in bacteria. Curr.Opin.Microbiol. **3**: 149-153.

Hernandez,N., and A.M.Weiner. 1986. Formation of the 3' end of U1 snRNA requires compatible snRNA promoter elements. Cell **47**: 249-258.

Hirose,Y., and J.L.Manley. 1998b. RNA polymerase II is an essential mRNA polyadenylation factor. Nature **395**: 93-96.

Hirose,Y., and J.L.Manley. 1998a. RNA polymerase II is an essential mRNA polyadenylation factor. Nature **395**: 93-96.

Hodel,A.E., P.D.Gershon, X.Shi, and F.A.Quiocho. 1996b. The 1.85 A structure of vaccinia protein VP39: a bifunctional enzyme that participates in the modification of both mRNA ends. Cell **85**: 247-256.

Hodel,A.E., P.D.Gershon, X.Shi, and F.A.Quiocho. 1996a. The 1.85 A structure of vaccinia protein VP39: a bifunctional enzyme that participates in the modification of both mRNA ends. Cell **85**: 247-256.

Howard,S.T., C.A.Ray, D.D.Patel, J.B.Antczak, and D.J.Pickup. 1999. A 43-nucleotide RNA cis-acting element governs the site-specific formation of the 3' end of a poxvirus late mRNA. Virology **255**: 190-204.

Huang,Y., and R.J.Maraia. 2001. Comparison of the RNA polymerase III transcription machinery in Schizosaccharomyces pombe, Saccharomyces cerevisiae and human. Nucleic Acids Res. **29**: 2675-2690.

Ink,B.S., and D.J.Pickup. 1990. Vaccinia virus directs the synthesis of early mRNAs containing 5' poly(A) sequences. Proc.Natl.Acad.Sci.U.S.A **87**: 1536-1540.

Jansa,P., C.Burek, E.E.Sander, and I.Grummt. 2001. The transcript release factor PTRF augments ribosomal gene transcription by facilitating reinitiation of RNA polymerase I. Nucleic Acids Res. **29**: 423-429.

Jansa,P., and I.Grummt. 1999b. Mechanism of transcription termination: PTRF interacts with the largest subunit of RNA polymerase I and dissociates paused transcription complexes from yeast and mouse. Mol.Gen.Genet. **262**: 508-514.

Jansa,P., and I.Grummt. 1999a. Mechanism of transcription termination: PTRF interacts with the largest subunit of RNA polymerase I and dissociates paused transcription complexes from yeast and mouse. Mol.Gen.Genet. **262**: 508-514.

Jansa,P., S.W.Mason, U.Hoffmann-Rohrer, and I.Grummt. 1998. Cloning and functional characterization of PTRF, a novel protein which induces dissociation of paused ternary transcription complexes. EMBO J. **17**: 2855-2864.

Jeong,S.W., W.H.Lang, and R.H.Reeder. 1995. The release element of the yeast polymerase I transcription terminator can function independently of Reb1p. Mol.Cell Biol. **15**: 5929-5936.

Jeong,S.W., W.H.Lang, and R.H.Reeder. 1996. The yeast transcription terminator for RNA polymerase I is designed to prevent polymerase slippage. J.Biol.Chem. **271**: 16104-16110.

Kates,J., and J.Beeson. 1970a. Ribonucleic acid synthesis in vaccinia virus. I. The mechanism of synthesis and release of RNA in vaccinia cores. J.Mol.Biol. **50**: 1-18.

Kates,J., and J.Beeson. 1970b. Ribonucleic acid synthesis in vaccinia virus. II. Synthesis of polyriboadenylic acid. J.Mol.Biol. **50**: 19-33.

Kates,J.R., and B.R.McAuslan. 1967. Poxvirus DNA-dependent RNA polymerase. Proc.Natl.Acad.Sci.U.S.A **58**: 134-141.

Katz,E. 1987. Thiosemicarbazones: inhibition of the growth of pox viruses and requirement for the growth of an isatin-beta-thiosemicarbazone dependent mutant. Rev.Clin.Basic Pharm. **6**: 119-130.

Keck,J.G., C.J.Baldick, Jr., and B.Moss. 1990. Role of DNA replication in vaccinia virus gene expression: a naked template is required for transcription of three late trans-activator genes. Cell **61**: 801-809.

Kim,D.E., and S.S.Patel. 2001. The kinetic pathway of RNA binding to the Escherichia coli transcription termination factor Rho. J.Biol.Chem. **276**: 13902-13910.

Korzheva,N., and A.Mustaev. 2001. Transcription elongation complex: structure and function. Curr.Opin.Microbiol. **4**: 119-125.

Kovacs,G.R., and B.Moss. 1996. The vaccinia virus H5R gene encodes late gene transcription factor 4: purification, cloning, and overexpression. J.Virol. **70**: 6796-6802.

Kuge,H., G.G.Brownlee, P.D.Gershon, and J.D.Richter. 1998. Cap ribose methylation of c-mos mRNA stimulates translation and oocyte maturation in Xenopus laevis. Nucleic Acids Res. **26**: 3208-3214.

Kuhn,A., I.Bartsch, and I.Grummt. 1990. Specific interaction of the murine transcription termination factor TTF I with class-I RNA polymerases. Nature **344**: 559-562.

Kulish,D., J.Lee, I.Lomakin, B.Nowicka, A.Das, S.Darst, K.Normet, and S.Borukhov. 2000. The functional role of basic patch, a structural element of Escherichia coli transcript cleavage factors GreA and GreB. J.Biol.Chem. **275**: 12789-12798.

Labhart,P. 1997. Transcript cleavage in an RNA polymerase I elongation complex. Evidence for a dissociable activity similar to but distinct from TFIIS. J.Biol.Chem. **272**: 9055-9061.

Lackner,C.A., and R.C.Condit. 2000. Vaccinia virus gene A18R DNA helicase is a transcript release factor. J.Biol.Chem. **275**: 1485-1494.

Lang,W.H., B.E.Morrow, Q.Ju, J.R.Warner, and R.H.Reeder. 1994. A model for transcription termination by RNA polymerase I. Cell **79**: 527-534.

Lang,W.H., T.Platt, and R.H.Reeder. 1998. Escherichia coli rho factor induces release of yeast RNA polymerase II but not polymerase I or III. Proc.Natl.Acad.Sci.U.S.A **95**: 4900-4905.

Lang,W.H., and R.H.Reeder. 1995. Transcription termination of RNA polymerase I due to a T-rich element interacting with Reb1p. Proc.Natl.Acad.Sci.U.S.A **92**: 9781-9785.

Latner,D.R., Y.Xiang, J.I.Lewis, J.Condit, and R.C.Condit. 2000. The vaccinia virus bifunctional gene J3 (nucleoside-2'-O-)- methyltransferase and poly(A) polymerase stimulatory factor is implicated as a positive transcription elongation factor by two genetic approaches. Virology **269**: 345-355.

LeRoy,G., G.Orphanides, W.S.Lane, and D.Reinberg. 1998. Requirement of RSF and FACT for transcription of chromatin templates in vitro. Science **282**: 1900-1904.

Lesnik,E.A., R.Sampath, H.B.Levene, T.J.Henderson, J.A.McNeil, and D.J.Ecker. 2001. Prediction of rho-independent transcriptional terminators in Escherichia coli. Nucleic Acids Res. **29**: 3583-3594.

Lewin,B. *Genes V*,. 1994. Oxford University Press, Oxford.

Li,J., and S.S.Broyles. 1993a. Recruitment of vaccinia virus RNA polymerase to an early gene promoter by the viral early transcription factor. J.Biol.Chem. **268**: 2773-2780.

Li,J., and S.S.Broyles. 1993b. The DNA-dependent ATPase activity of vaccinia virus early gene transcription factor is essential for its transcription activation function. J.Biol.Chem. **268**: 20016-20021.

Lin-Marq,N., and S.G.Clarkson. 1998. Efficient synthesis, termination and release of RNA polymerase III transcripts in Xenopus extracts depleted of La protein. EMBO J. **17**: 2033-2041.

Lodish,H.F., and J.E.Darnell. *Molecular cell biology*, Scientific American Books, New York.

Logan,J., E.Falck-Pedersen, J.E.Darnell, Jr., and T.Shenk. 1987. A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. Proc.Natl.Acad.Sci.U.S.A **84**: 8306-8310.

Mahr,A., and B.E.Roberts. 1984. Arrangement of late RNAs transcribed from a 7.1-kilobase EcoRI vaccinia virus DNA fragment. J.Virol. **49**: 510-520.

Maraia,R.J., and R.V.Intine. 2001. Recognition of nascent RNA by the human La antigen: conserved and divergent features of structure and function. Mol.Cell Biol. **21**: 367-379.

Maraia,R.J., D.J.Kenan, and J.D.Keene. 1994. Eukaryotic transcription termination factor La mediates transcript release and facilitates reinitiation by RNA polymerase III. Mol.Cell Biol. **14**: 2147-2158.

Martin,S.A., and B.Moss. 1975. Modification of RNA by mRNA guanylyltransferase and mRNA (guanine-7-)methyltransferase from vaccinia virions. J.Biol.Chem. **250**: 9330-9335.

Martin,S.A., and B.Moss. 1976. mRNA guanylyltransferase and mRNA (guanine-7-)-methyltransferase from vaccinia virions. Donor and acceptor substrate specificites. J.Biol.Chem. **251**: 7313-7321.

Mason,S.W., E.E.Sander, and I.Grummt. 1997. Identification of a transcript release activity acting on ternary transcription complexes containing murine RNA polymerase I. EMBO J. **16**: 163-172.

Massung,R.F., L.I.Liu, J.Qi, J.C.Knight, T.E.Yuran, A.R.Kerlavage, J.M.Parsons, J.C.Venter, and J.J.Esposito. 1994. Analysis of the complete genome of smallpox variola major virus strain Bangladesh-1975. Virology **201**: 215-240.

McCracken,S., N.Fong, K.Yankulov, S.Ballantyne, G.Pan, J.Greenblatt, S.D.Patterson, M.Wickens, and D.L.Bentley. 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature **385**: 357-361.

McCracken,S., E.Rosonina, N.Fong, M.Sikes, A.Beyer, K.O'Hare, S.Shuman, and D.Bentley. 1998. Role of RNA polymerase II carboxy-terminal domain in coordinating transcription with RNA processing. Cold Spring Harb.Symp.Quant.Biol. **63**: 301-309.

Meis,R.J., and R.C.Condit. 1991. Genetic and molecular biological characterization of a vaccinia virus gene which renders the virus dependent on isatin-beta-thiosemicarbazone (IBT). Virology **182**: 442-454.

Mohamed,M.R., D.R.Latner, R.C.Condit, and E.G.Niles. 2001. Interaction between the J3R subunit of vaccinia virus poly(A) polymerase and the H4L subunit of the viral RNA polymerase. Virology **280**: 143-152.

Mohamed,M.R., and E.G.Niles. 2000. Interaction between nucleoside triphosphate phosphohydrolase I and the H4L subunit of the viral RNA polymerase is required for vaccinia virus early gene transcript release. J.Biol.Chem. **275**: 25798-25804.

Mohamed,M.R., and E.G.Niles. 2001. The viral RNA polymerase H4L subunit is required for Vaccinia virus early gene transcription termination. J.Biol.Chem. **276**: 20758-20765.

Mooney,R.A., I.Artsimovitch, and R.Landick. 1998. Information processing by RNA polymerase: recognition of regulatory signals during RNA chain elongation. J.Bacteriol. **180**: 3265-3275.

Moreland,R.J., J.S.Hanas, J.W.Conaway, and R.C.Conaway. 1998. Mechanism of action of RNA polymerase II elongation factor Elongin. Maximal stimulation of elongation requires conversion of the early elongation complex to an Elongin-activable form. J.Biol.Chem. **273**: 26610-26617.

Moreland,R.J., F.Tirode, Q.Yan, J.W.Conaway, J.M.Egly, and R.C.Conaway. 1999. A role for the TFIIH XPB DNA helicase in promoter escape by RNA polymerase II. J.Biol.Chem. **274**: 22127-22130.

Moss,B. 1990a. Regulation of orthopoxvirus gene expression. Curr.Top.Microbiol.Immunol. **163**: 41-70.

Moss,B. 1990b. Regulation of vaccinia virus transcription. Annu.Rev.Biochem. **59**: 661-688.

Moss,B. 1996. Poxviridae: The Viruses and Their Replication. In *Fundamental Virology* (ed. B.N.Fields, D.M.Knipe, and P.M.Howley), pp. 1163-1197. Lippincott-Raven, Philadelphia.

Moss,B., E.N.Rosenblum, and A.Gershowitz. 1975. Characterization of a polyriboadenylate polymerase from vaccinia virions. J.Biol.Chem. **250**: 4722-4729.

Nevins,J.R., and W.K.Joklik. 1977. Isolation and partial characterization of the poly(A) polymerases from HeLa cells infected with vaccinia virus. J.Biol.Chem. **252**: 6939-6947.

Nudler,E. 1999. Transcription elongation: structural basis and mechanisms. J.Mol.Biol. **288**: 1-12.

Orphanides,G., G.LeRoy, C.H.Chang, D.S.Luse, and D.Reinberg. 1998. FACT, a factor that facilitates transcript elongation through nucleosomes. Cell **92**: 105-116.

Osheim,Y.N., N.J.Proudfoot, and A.L.Beyer. 1999. EM visualization of transcription by RNA polymerase II: downstream termination requires a poly(A) signal but not transcript cleavage. Mol.Cell **3**: 379-387.

Pacha,R.F., and R.C.Condit. 1985. Characterization of a temperature-sensitive mutant of vaccinia virus reveals a novel function that prevents virus-induced breakdown of RNA. J.Virol. **56**: 395-403.

Pacha,R.F., R.J.Meis, and R.C.Condit. 1990. Structure and expression of the vaccinia virus gene which prevents virus-induced breakdown of RNA. J.Virol. **64**: 3853-3863.

Platt,T. 1994. Rho and RNA: models for recognition and response. Mol.Microbiol. **11**: 983-990.

Price,D.H. 2000. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. Mol.Cell Biol. **20**: 2629-2634.

Proudfoot,N.J. 1989b. How RNA polymerase II terminates transcription in higher eukaryotes. Trends Biochem.Sci. **14**: 105-110.

Proudfoot,N.J. 1989a. How RNA polymerase II terminates transcription in higher eukaryotes. Trends Biochem.Sci. **14**: 105-110.

Reeder,R.H. and W.H.Lang. 1997. Terminating transcription in eukaryotes: lessons learned from RNA polymerase I. Trends Biochem.Sci. **22**: 473-477.

Richardson,J.P. 1993. Transcription termination. Crit Rev.Biochem.Mol.Biol. **28**: 1-30.

Richardson,L.V. and J.P.Richardson. 1996. Rho-dependent termination of transcription is governed primarily by the upstream Rho utilization (rut) sequences of a terminator. J.Biol.Chem. **271**: 21597-21603.

Roberts,J.W. 1969. Termination factor for RNA synthesis. Nature **224**: 1168-1174.

Roberts,J.W. 1975. Transcription termination and late control in phage lambda. Proc.Natl.Acad.Sci.U.S.A **72**: 3300-3304.

Roberts,J.W. 1988. Phage lambda and the regulation of transcription termination. Cell **52**: 5-6.

Rosales,R., N.Harris, B.Y.Ahn, and B.Moss. 1994. Purification and identification of a vaccinia virus-encoded intermediate stage promoter-specific transcription factor that has homology to eukaryotic transcription factor SII (TFIIS) and an additional role as a viral RNA polymerase subunit. J.Biol.Chem. **269**: 14260-14267.

Rosales,R., G.Sutter, and B.Moss. 1994. A cellular factor is required for transcription of vaccinia viral intermediate-stage genes. Proc.Natl.Acad.Sci.U.S.A **91**: 3794-3798.

Schnapp,G., A.Schnapp, H.Rosenbauer, and I.Grummt. 1994. TIF-IC, a factor involved in both transcription initiation and elongation of RNA polymerase I. EMBO J. **13**: 4028-4035.

Schnierle,B.S., P.D.Gershon, and B.Moss. 1992. Cap-specific mRNA (nucleoside-O2'-)-methyltransferase and poly(A) polymerase stimulatory activities of vaccinia virus are mediated by a single protein. Proc.Natl.Acad.Sci.U.S.A **89**: 2897-2901.

Schnierle,B.S., P.D.Gershon, and B.Moss. 1994. Mutational analysis of a multifunctional protein, with mRNA 5' cap- specific (nucleoside-2'-O-)-methyltransferase and 3'-adenylyltransferase stimulatory activities, encoded by vaccinia virus. J.Biol.Chem. **269**: 20700-20706.

Selby,C.P., and A.Sancar. 1997. Cockayne syndrome group B protein enhances elongation by RNA polymerase II. Proc.Natl.Acad.Sci.U.S.A **94**: 11205-11209.

Senkevich,T.G., E.V.Koonin, J.J.Bugert, G.Darai, and B.Moss. 1997. The genome of molluscum contagiosum virus: analysis and comparison with other poxviruses. Virology **233**: 19-42.

Shchelkunov,S.N., A.V.Totmenin, V.N.Loparev, P.F.Safronov, V.V.Gutorov, V.E.Chizhikov, J.C.Knight, J.M.Parsons, R.F.Massung, and J.J.Esposito. 2000. Alastrim smallpox variola minor virus genome DNA sequences. Virology **266**: 361-386.

Shi,X., T.G.Bernhardt, S.M.Wang, and P.D.Gershon. 1997. The surface region of the bifunctional vaccinia RNA modifying protein VP39 that interfaces with Poly(A) polymerase is remote from the RNA binding cleft used for its mRNA 5' cap methylation function. J.Biol.Chem. **272**: 23292-23302.

Shi,X., P.Yao, T.Jose, and P.Gershon. 1996. Methyltransferase-specific domains within VP-39, a bifunctional protein that participates in the modification of both mRNA ends. RNA. **2**: 88-101.

Shilatifard,A. 1998a. Factors regulating the transcriptional elongation activity of RNA polymerase II. FASEB J. **12**: 1437-1446.

Shilatifard,A. 1998b. Identification and purification of the Holo-ELL complex. Evidence for the presence of ELL-associated proteins that suppress the transcriptional inhibitory activity of ELL. J.Biol.Chem. **273**: 11212-11217.

Shuman,S., S.S.Broyles, and B.Moss. 1987. Purification and characterization of a transcription termination factor from vaccinia virions. J.Biol.Chem. **262**: 12372-12380.

Shuman,S. and B.Moss. 1990. Purification and use of vaccinia virus messenger RNA capping enzyme. Methods Enzymol. **181**: 170-180.

Simpson,D.A., and R.C.Condit. 1994. The vaccinia virus A18R protein plays a role in viral transcription during both the early and the late phases of infection. J.Virol. **68**: 3642-3649.

Simpson,D.A., and R.C.Condit. 1995. Vaccinia virus gene A18R encodes an essential DNA helicase. J.Virol. **69**: 6131-6139.

Stitt,B.L. 2001. Escherichia coli transcription termination factor Rho binds and hydrolyzes ATP using a single class of three sites. Biochemistry **40**: 2276-2281.

Sudarsanam,P., and F.Winston. 2000. The Swi/Snf family nucleosome-remodeling complexes and transcriptional control. Trends Genet. **16**: 345-351.

Thompson,C.L., and R.C.Condit. 1986. Marker rescue mapping of vaccinia virus temperature-sensitive mutants using overlapping cosmid clones representing the entire virus genome. Virology **150**: 10-20.

Tschochner,H. 1996. A novel RNA polymerase I-dependent RNase activity that shortens nascent transcripts from the 3' end. Proc.Natl.Acad.Sci.U.S.A **93**: 12914-12919.

Vagner,S., C.Vagner, and I.W.Mattaj. 2000. The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. Genes Dev. **14**: 403-413.

van Hoof,A., and R.Parker. 1999. The exosome: a proteasome for RNA? Cell **99**: 347-350.

Vasserot,A.P., F.J.Schaufele, and M.L.Birnstiel. 1989. Conserved terminal hairpin sequences of histone mRNA precursors are not involved in duplex formation with the U7 RNA but act as a target site for a distinct processing factor. Proc.Natl.Acad.Sci.U.S.A **86**: 4345-4349.

von Hippel,P.H. 1998. An integrated model of the transcription complex in elongation, termination, and editing. Science **281**: 660-665.

Vos,J.C., M.Sasker, and H.G.Stunnenberg. 1991. Vaccinia virus capping enzyme is a transcription initiation factor. EMBO J. **10**: 2553-2558.

Vos,J.C., and H.G.Stunnenberg. 1988. Derepression of a novel class of vaccinia virus genes upon DNA replication. EMBO J. **7**: 3487-3492.

Wada,T., T.Takagi, Y.Yamaguchi, A.Ferdous, T.Imai, S.Hirose, S.Sugimoto, K.Yano, G.A.Hartzog, F.Winston, S.Buratowski, and H.Handa. 1998. DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev. **12**: 343-356.

Wang,Z., L.Bai, Y.J.Hsieh, and R.G.Roeder. 2000. Nuclear factor 1 (NF1) affects accurate termination and multiple-round transcription by human RNA polymerase III. EMBO J. **19**: 6823-6832.

Wang,Z., and R.G.Roeder. 1998. DNA topoisomerase I and PC4 can interact with human TFIIIC to promote both accurate termination and transcription reinitiation by RNA polymerase III. Mol.Cell **1**: 749-757.

Wei,C.M., and B.Moss. 1975. Methylated nucleotides block 5'-terminus of vaccinia virus messenger RNA. Proc.Natl.Acad.Sci.U.S.A **72**: 318-322.

Weisberg,R.A., and M.E.Gottesman. 1999. Processive antitermination. J.Bacteriol. **181**: 359-367.

Willer,D.O., G.McFadden, and D.H.Evans. 1999. The complete genome sequence of shope (rabbit) fibroma virus. Virology **264**: 319-343.

Wind,M., and D.Reines. 2000. Transcription elongation factor SII. Bioessays **22**: 327-336.

Woodson,B. and W.K.Joklik. 1965. The inhibition of vaccinia virus multiplication by isatin-beta- thiosemicarbazone. Proc.Natl.Acad.Sci.U.S.A **54**: 946-953.

World Resource Institute. Biodiversity - How many species are there? World Resource Institute . 2001. http://www.wri.org/biodiv/b02-gbs.html

Wright,C.F., and A.M.Coroneos. 1995. The H4 subunit of vaccinia virus RNA polymerase is not required for transcription initiation at a viral late promoter. J.Virol. **69**: 2602-2604.

Wright,C.F., A.E.Hubbs, S.K.Gunasinghe, and B.W.Oswald. 1998. A vaccinia virus late transcription factor copurifies with a factor that binds to a viral late promoter and is complemented by extracts from uninfected HeLa cells. J.Virol. **72**: 1446-1451.

Wu,S.Y. and T.Platt. 1993. Transcriptional arrest of yeast RNA polymerase II by Escherichia coli rho protein in vitro. Proc.Natl.Acad.Sci.U.S.A **90**: 6606-6610.

Xiang,Y., D.R.Latner, E.G.Niles, and R.C.Condit. 2000a. Transcription elongation activity of the vaccinia virus J3 protein in vivo is independent of poly(A) polymerase stimulation. Virology **269**: 356-369.

Xiang,Y., D.R.Latner, E.G.Niles, and R.C.Condit. 2000b. Transcription elongation activity of the vaccinia virus J3 protein in vivo is independent of poly(A) polymerase stimulation. Virology **269**: 356-369.

Xiang,Y., D.A.Simpson, J.Spiegel, A.Zhou, R.H.Silverman, and R.C.Condit. 1998. The vaccinia virus A18R DNA helicase is a postreplicative negative transcription elongation factor. J.Virol. **72**: 7012-7023.

Xie,Z. and D.Price. 1997. Drosophila factor 2, an RNA polymerase II transcript release factor, has DNA-dependent ATPase activity. J.Biol.Chem. **272**: 31902-31907.

Xie,Z. and D.H.Price. 1996. Purification of an RNA polymerase II transcript release factor from Drosophila. J.Biol.Chem. **271**: 11043-11046.

Xie,Z. and D.H.Price. 1998. Unusual nucleic acid binding properties of factor 2, an RNA polymerase II transcript release factor. J.Biol.Chem. **273**: 3771-3777.

Yamaguchi,Y., T.Takagi, T.Wada, K.Yano, A.Furuya, S.Sugimoto, J.Hasegawa, and H.Handa. 1999. NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. Cell **97**: 41-51.

Yan,Q., R.J.Moreland, J.W.Conaway, and R.C.Conaway. 1999. Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. J.Biol.Chem. **274**: 35668-35675.

Yonaha,M. and N.J.Proudfoot. 1999. Specific transcriptional pausing activates polyadenylation in a coupled in vitro system. Mol.Cell **3**: 593-600.

Yuen,L. and B.Moss. 1987. Oligonucleotide sequence signaling transcriptional termination of vaccinia virus early genes. Proc.Natl.Acad.Sci.U.S.A **84**: 6417-6421.

Zeng,C. and S.M.Berget. 2000. Participation of the C-terminal domain of RNA polymerase II in exon definition during pre-mRNA splicing. Mol.Cell Biol. **20**: 8290-8301.

Zhao,J., L.Hyman, and C.Moore. 1999. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol.Mol.Biol.Rev. **63**: 405-445.

BIOGRAPHICAL SKETCH

Donald R. Latner II was born in east Tennessee and was instilled from the beginning with a strong sense of "Appalachian values" (let the reader interpret). Learning from a large and close knit family, he developed a love of God and creation that ultimately led him to study the very fiber of life itself. He spent his first year of college at East Tennessee State University and finished his bachelor's degree in genetics at the University of Georgia in 1996. He completed his Ph.D. at the University of Florida in 2001 as a member of the first class of the Interdisciplinary Program in Biomedical Sciences (IDP).

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Richard C. Condit, Chair
Professor of Molecular Genetics and Microbiology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Arthur Edison
Assistant Professor of Biochemistry and Molecular Biology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

James Bert Flanegan
Professor of Biochemistry and Molecular Biology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Sue Moyer
Professor of Molecular Genetics and Microbiology

This dissertation was submitted to the Graduate Faculty of the College of Medicine and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 2001

Wayne I McCormack

Dean, College of Medicine

Winfred M Phillips

Dean, Graduate School